An investigation of heterologous expression of human steroidogenic cytochromes P450 in yeasts

Norbert Wilhelm Kolar

Dissertation
presented for the
Degree of Doctor of Philosophy
(Biochemistry)
in the
Faculty of Science
at the
University of Stellenbosch

Promoter: Prof. P. Swart
Co-promoter: Dr. A. C. Swart
Department of Biochemistry, University of Stellenbosch, Stellenbosch, South Africa.
April 2005

Stellenbosch University http://scholar.sun.ac.za

Declaration:

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature:.

Date:

Stellenbosch University http://scholar.sun.ac.za

SUMMARY

This study:

1. Compares various heterologous expression systems for high-level expression of cytochromes P450. Limitations of the existing cytochromes P450 expression systems are discussed and possibilities to improve the expression yields of human steroidogenic enzymes, are suggested. In addition the potential applications of human steroidogenic cytochromes P450 expressed in *Pichia pastoris* are illustrated.

2. Describes the cloning and extracellular expression of a recombinant full-length human cytochrome P450 17α-hydroxylase (P45017α) enzyme in *Saccharomyces cerevisiae*. After the optimisation of expression conditions, it was shown that this system is not suitable for the expression of full-length human P45017α.

3. Describes the cloning and extracellular expression of the full-length human cytochromes P45017α, aromatase, b_5 and truncated human cytochrome P45017α in *P. pastoris*. The limitations using *P. pastoris* as an export system for expressed P450 enzymes were pointed out.

4. Describes the cloning and intracellular expression of the full-length human cytochrome P45017α in *P. pastoris* as well as the functional expression of human P45017α in *P. pastoris*, showing progesterone conversion to 17α-hydroxyprogesterone and 16α-hydroxyprogesterone *in vivo,* for the first time.

5. Evaluates developed methods for the preparation of microsomes from *P. pastoris* expressing human P45017α and the spectral characterisation of detergent solubilised human P45017α.

6. Describes the development of protocols for the purification of human P45017α from *P. pastoris* microsomes.

Stellenbosch University http://scholar.sun.ac.za

OPSOMMING

Hierdie ondersoek:

1. Vergelyk verskillende heterologiese proteïen uitdrukkings-sisteme vir die preperatiewe produksie van sitochrome P450. Die tekortkomings van bestaande sitochroom P450-uitdrukkings-sisteme word bespreek en moontlikhede om die opbrengs van menslike steroïedogeniese ensieme te verbeter word voorgestel. Die potensiële aanwendings van menslike steroïedogeniese sitochrome P450, wat in *Pichia pastoris* uitgedruk word, word ook geïllustreer.

2. Beskryf die klonering en ekstrasellulêre uitdrukking van die rekombinante vollengte menslike sitochroom P45017α-hidroksilase (P45017α) ensiem in *Saccharomyces cerevisiae*. Na optimisering van die kondisies vir die uitdrukking kon aangetoon word dat hierdie sisteem nie geskik is vir die uitdruk en sekresie van vollengte menslike P45017α nie.

3. Beskryf die klonering en ekstrasellulêre uitdrukking van die vollengte menslike sitochrome P45017α, aromatase, b5 en verkorte menslike sitochroom P45017α in *P. pastoris*. Die beperkinge van *P. pastoris* as 'n uitvoersisteem vir die uitdrukking van P450 ensieme word bespreek.

4. Besryf die klonering en intrasellulêre ekspressie van die vollengte menslike sitochroom P450 17α. Die funksionele ekspressie van menslike sitochroom P450 17α in *P. pastoris* is vir die eerste keer gekarakteriseer.

5. Evalueer die ontwikkelde metodes vir die voorbereiding van mikrosome van *P. pastoris* wat menslike P45017α uitdruk en karakteriseer die detergent opgelosbare menslike P45017α t.o.v. spektroskopiese eienskappe.

6. Beskryf die ontwikkelling van protokolle vir die suiwering van die uitgedrukte menslike P45017α vanuit *P. pastoris* mikrosome.

Stellenbosch University https://scholar.sun.ac.za

ZUSAMMENFASSUNG

Diese Studie:

1. Vergleicht verschiedene heterologe Expressionssystemen für die Überproduktion von Cytochrom P450. Die Begrenzungen von den existierenden Cytochrom P450-Expressionssystemen wurden besprochen und Möglichkeiten um die Produktion von menschlichen Steroid-Enzyme zu verbessern wurden vorgeschlagen. Die potenzialen Anwendungen von menschlichen Cytochrom P450 Steroid-Enzymen, welche vom Hefe *Pichia pastoris* produziert werden, wurden illustriert.

2. Beschreibt das Klonieren und extrazelluläre Expression von dem rekominanten menschlichen Cytochrom 17α-hydroxylase (P45017α) Enzym in *Saccharomyces cerevisiae.* Nachdem der Optimierung der Expressionskonditionen konnte gezeigt werden, daß das Expressionssystem für die Produktion von dem menschlichen P45017α nicht geeignet ist.

3. Beschreibt das Klonieren und extrazelluläre Expression von dem rekominanten menschlichen Cytochromen P45017α, aromatase, b_5 und ein verkürztes menschliches Cytochrom P450 in *P. pastoris*. Das Benutzen von *P. pastoris* als ein extrazelluläres Expressionssystem hat Begrenzungen die besprochen werden.

4. Beschreibt das Klonieren und intrazelluläre Expression von dem rekominanten menschlichen Cytochrom P45017α. Die funktionelle Expression von dem menschlichen Cytochrom P45017α in *P. pastoris* wurde für das erste Mal karakterisiert.

5. Evaluiert entwickelte Metoden für die Vorbereitung von Mikrosomen von *P. pastoris* die das menschliche P45017α heterolog exprimieren und die spektroskopischen Eigenschaften von dem detergens auflösbare menschliche P45017α karakterisiert.

6. Beschreibt die Entwicklung von Protokollen für die Säuberung von dem P45017α Enzym das in den Mikrosomen von *P. pastoris* vorkommt.

To my **parents** for giving me life
and to **Gisela** for sharing it with me.

Stellenbosch University http://scholar.sun.ac.za

ACKNOWLEDGEMENTS

I herewith wish to express my sincerest gratitude to the following persons:

Prof. P. Swart for giving me the freedom to develop an understanding of science independently and for the stimulating discussions

Dr. A. C. Swart for the invaluable lessons in molecular biology you so willingly gave

Ralie for always supplying me with the necessary chemicals

Brett for walking this long road to freedom with me

My parents for their sacrifice in order to provide me with the educational basis without which this achievement would not be possible

Claudia for fulfilling the role of our parents in a difficult time for us both and giving all your love and support

George and Bernice for adopting me as your own and not allowing me to give up on my dreams

Gisela for marrying me and helping to carry some of the burden during the stressful times of this degree.

Stellenbosch University http://scholar.sun.ac.za

TABLE OF CONTENTS

PART I: INTRODUCTION, LITERATURE REVIEWS AND GOAL OF THE STUDY

CHAPTER 1: Introduction 5

CHAPTER 2: Literature reviews 8

SECTION 2.1: An insight in structure, function and role of two key steroidogenic enzymes P45017α and aromatase 8

SECTION 2.2: Comparative study of heterologous expression systems for high-level expression of cytochromes P450 — Potential Applications of human steroidogenic cytochromes P450 expressed in *Pichia pastoris* 25

SECTION 2.3: Heterologous protein secretion by yeasts' secretory pathway 61

CHAPTER 3: Goals of the study 72

References 73

PART II: CLONING AND EXTRACELLULAR EXPRESSION OF RECOMBINANT HUMAN CYTOCHROME P45017α (17α-HYDROXYLASE/17,20-LYASE) IN *SACCHAROMYCES CEREVISIAE*

Introduction 95

CHAPTER 1: Cloning and construction of the full-length human CYP17 into YEpFLAG expression vector and transformation in *S. cerevisiae* 98

CHAPTER 2: Culture of *S. cerevisiae* expressing P45017α and optimisation of protein expression 108

CHAPTER 3: Discussion 119

References 125

PART III: CLONING AND EXTRACELLULAR EXPRESSION OF THE FULL-LENGTH HUMAN CYTOCHROMES P45017α, AROMATASE, b5 AND TRUNCATED HUMAN CYTOCHROME P45017α IN *PICHIA PASTORIS*

Introduction 132

CHAPTER 1: **Plasmid preparations, transformation and screening of *P. pastoris*** 137

1.1: Cloning and construction of the full-length human CYP17 into pPIC9K expression vector and integration in *P. pastoris* 137

1.2: Cloning and construction of the truncated human CYP17 (truncCYP17) into the pPIC9K expression vector 152

1.3: Cloning and construction of the full-length human b_5 gene (CYB5) into the pPIC9K expression vector 159

1.4: Cloning and construction of full-length human CYP19 into the pPIC9K expression vector 168

CHAPTER 2: **Culture of *P. pastoris* expressing different proteins** 177

CHAPTER 3: **Discussion** 188

References 193

PART IV: CLONING, INTRACELLULAR EXPRESSION AND PURIFICATION OF RECOMBINANT HUMAN CYTOCHROME P45017α (17α-HYDROXYLASE/17,20-LYASE) IN *PICHIA PASTORIS*

Introduction 202

CHAPTER 1: **Construction of the full-length human CYP17 into pPIC3.5K expression vector and integration in *P. pastoris*** 203

CHAPTER 2: **Culture of *P. pastoris* expressing P45017α and successful expression of the protein** 211

CHAPTER 3: **Development and evaluation of methods for the preparation of microsomes from *P. pastoris* expressing human P45017α** 217

CHAPTER 4: **The development and evaluation of a protocol for the purification of P45017α from microsomes** 228

CHAPTER 5: **Discussion** 243

References 250

CONCLUSION 253

APPENDIX A 254

APPENDIX B 263

ABBREVIATIONS

17β-HSD	17β-hydroxysteroid dehydrogenase
3β-HSD	3β-hydroxysteroid dehydrogenase/Δ5-Δ4 isomerase
17-KSR	17-ketosteroid reductase
aa	Amino acid
aromatase	Cytochrome P450 aromatase
b_5	Cytochrome b_5
BAP	Bacterial alkaline phosphatase
bp	Base pair
CO	Carbon monoxide
cpm	Counts per million
CYB5	Cytochrome b_5 gene
CYP17	Cytochrome P450 17α-hydroxylase gene
CYP19	Cytochrome P450 19α-hyroxylase gene
Da	Daltons
DEPC	Diethylpyrocarbonate
DHEA	Dehydroepiandrosterone
DHEA-S	Dehydroepiandrosterone sulphate
DIG	Dioxigenin
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
GFP	Green fluorescent protein
Glc	Glucose
GlcNAc	N-acetylglucosamine
IB	Inclusion bodies
kb	Kilo bases
Kex1p	Carboxypeptidase B
Kex2p	Endoprotease
lacZ gene	β-galactosidase gene
M	Mannose
MCO	Molecular cut-off
MFα1	Prepro-α-factor
NAC	Nascent-polypeptide-associated complex
$NADP^+$	Nicotinamide adenine dinucleotide phosphate (oxidised)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
OR	NADPH-P450 oxidoreductase
OST	Oligosaccharyltransferase complex
P450aldo	Cytochrome P450 aldosterone synthase
P450arom	Cytochrome P450 aromatase
P450c11	Cytochrome P450 11β-hydroxylase
P45017α	Cytochrome P450 17α-hydroxylase
P450c21	Cytochrome P450 21-hydroxylase

Stellenbosch University http://scholar.sun.ac.za

P450scc	Cytochrome P450 side chain cleavage
PCR	Polymerase chain reaction
PDI	Protein disulfide isomerase
PHO2	Acid phosphatase
PMSF	Phenyl methyl sulfonyl fluoride
SA-I	Type I signal-anchor
SA-II	Type II signal-anchor
SAS	Signal anchor sequence
SC^{-TRP}	Synthetic complete medium without tryptophane
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
Sec4p	GTPase protein complex
SNARE	Soluble N-ethylmaleinide sensitive factor attachment protein receptor
SRP	Signal recognition particle
Ste13	Dipeptidyl aminopeptidase A
SUC	Yeast invertase
WT	Wild-type strain

Stellenbosch University http://scholar.sun.ac.za

PART I

INTRODUCTION, LITERATURE REVIEWS AND GOAL OF THE STUDY

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 1: INTRODUCTION 5

CHAPTER 2: LITERATURE REVIEWS 8

SECTION 2.1: An insight in structure, function and role of two key steroidogenic enzymes P45017α and aromatase 8

2.1.1 Cytochromes P450 8

2.1.1.1 The mechanism of cytochromes P450 9

2.1.1.2 Membrane topology of cytochromes P450 together with their electron donors 11

2.1.1.2.1 Microsomal P450 monooxygenases 12

2.1.1.2.1.1 Mechanism of action of cytochromes P450 and their electron carriers 12

2.1.1.2.1.2 Mechanism of membrane insertion and protein retention 12

2.1.2 Enzymes in the steroidogenic pathways and steroid hydroxylases 14

2.1.2.1 Role and mechanism of P45017α in gonadal and adrenal steroidogenesis 17

2.1.2.1.1 Role of the 17α-hydroxylase and C17,20-lyase reaction 17

2.1.2.1.2 Mechanism of 17α-hydroxylase 18

2.1.2.1.3 Mechanism and regulation of 17,20-lyase activity 18

2.1.2.2 Role and mechanism of aromatase in gonadal steroidogenesis 20

2.1.2.2.1 Reaction mechanism of aromatase 21

2.1.3 Sequence and structural similarities of P45017α and aromatase genes 21

SECTION 2.2: Comparative study of heterologous expression systems for high-level expression of cytochromes P450 — Potential Applications of human steroidogenic cytochromes P450 expressed in *Pichia pastoris* 25

2.2.1 Introduction 25

2.2.2 Factors influencing high P450 expression levels in heterologous expression systems 28

2.2.2.1 The host 28

2.2.2.2 Co-expression of cytochromes P450 with OR 29

2.2.2.3 Co-expression of P450 enzymes with chaperones 33

2.2.2.4 Transcriptional control: vectors and promoters 33

Stellenbosch University http://scholar.sun.ac.za

2.2.2.5 Translational efficiency: strong initiation sites, secondary structures and codon usage 37
2.2.2.6 Fermentation conditions of high cell density growth in fermentor cultures 39
2.2.2.7 Fusion of cytochromes P450 to signal peptides and secretion 42
2.2.3 Genetically-engineered human steroidogenic cytochromes P450 from P. pastoris - potential applications 43
2.2.3.1 Potential applications using pure proteins 43
2.2.3.1.1 Analysing structures 43
2.2.3.1.2 Determination of substrate specificity 45
2.2.3.1.3 Elucidating electron transfer of redox partner 45
2.2.3.1.4 Production of antibodies 46
2.2.3.2 Potential applications using whole cells 47
2.2.3.2.1 Regulation of steroid metabolism 47
2.2.3.2.2 Testing mutagenicity and carcinogenicity 48
2.2.3.2.3 Manufacturing of metabolite in bioreactors 48
2.2.4 Conclusion 48

Table 2.2 A **51**

Table 2.2 B **56**

Table 2.2 C **57**

Table 2.2 D **60**

SECTION 2.3: Heterologous protein secretion by yeasts' secretory pathway **61**
2.3.1 Introduction 61
2.3.1.1 Passage of protein from the cytoplasm to ER and processing in the ER 62
2.3.1.1.1 Leader sequences for initiation of secretion 63
2.3.1.1.2 The translocation of the protein from the cytoplasm to the ER 65
2.3.1.1.3 The processing of the prepro-α-factor during translocation 66
2.3.1.1.4 Correct folding of heterologous proteins and N- glycosylation 66
2.3.1.2 Transport from the ER to the Golgi complex 67
2.3.1.3 Golgi processing and passage from Golgi to secretory vesicles 68
2.3.1.3.1 Outer chain glycosylation 68
2.3.1.3.2 Proteolytic processing step 69

Stellenbosch University http://scholar.sun.ac.za

2.3.1.4 Passage of the Golgi complex-derived vesicles and fusion to the plasma membrane 69

2.3.1.5 Passage of proteins through cell wall 70

2.3.2 Discussion 71

CHAPTER 3: GOALS OF THE STUDY **72**

REFERENCES **73**

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 1: INTRODUCTION

The cytochrome P450 dependent enzymes, a diverse group of heme-containing enzymes found in all prokaryotic and eukaryotic species, were discovered in the 1940s, during a series of experimental reactions involving molecular oxygen. These hemoproteins are defined as monooxygenases, because they catalyse the incorporation of a single oxygen atom from molecular oxygen into a large variety of endogenous and foreign lipophilic compounds. These reactions proved to be important in the metabolism of numerous groups of xenobiotics[1], including pesticides, drugs and hydrocarbons (1, 2). Not only are the P450-dependent enzymes essential in detoxification systems, but they also play an important role in the metabolism of endogenous substances produced within the body (3). These substances can be fat-soluble vitamins, fatty acids, eicosanoids, alkaloids and most importantly steroids.

Steroids are a unique group of hormones responsible for stress management, the maintenance of water and mineral balance and control of male and female sex characteristics in mammals. These steroids are synthesised by a unique set of enzymes, the steroidogenic cytochromes, which are localised in tissue-specific environments such as placenta, adrenals and gonads. Three of these essential enzymes, cytochrome P450 17α-hydroxylase (P45017α), cytochrome b_5 (b_5) and cytochrome P450arom (aromatase) are the main focus of this study and an overview of their structure, function and role in steroidogenesis is given in **Part I Section 2.1**.

Pure proteins are required to better understand the mechanism and structure/function relationship of the steroidogenic P450 enzymes. The low concentrations and legal implications involved in obtaining the P450 enzymes from human tissue for the direct isolation and purification lead to the development of heterologous expression systems. The heterologous expression systems for cytochrome P450-dependent enzymes, reviewed in **Part I Section 2.2,** has become an attractive tool in P450 biotechnology for the expression of cDNA encoding specific P450 enzymes and also offers several advantages in the research fields, pharmacology and toxicology. An ultrapure protein

---

[1] Chemical substances that are foreign to the biological system. They include naturally occurring compounds, drugs, environmental agents, carcinogens, insecticides, etc.

Stellenbosch University http://scholar.sun.ac.za

can, for instance, be produced to raise specific antibodies, which could serve as a diagnostic tool (i.e. immunohistochemistry) or for immunoinhibition studies. The potential to acquire large quantities of the native enzyme via heterologous expression, will facilitate the crystallisation of these extremely labile membrane proteins. In addition, heterologous expression systems can also help to investigate the functional relevance of genetic polymorphisms of P450 enzymes. By altering certain residues with site-directed mutagenesis, followed by heterologous expression, the structure/function relationship associated with cytochromes P450 can be elucidated.

A comprehensive updated study of heterologous expression systems for cytochromes P450, used in the past 20 years, is tabulated in **Part I, Section 2.2**. This compilation shows that most of the expression systems did not produce large quantities of P450 enzymes. In **section 2.2** the limitations of the existing cytochrome P450 expression systems are discussed and possibilities to improve the expression yields of steroidogenic P450 enzymes are suggested.

The first strategy was to export the expressed P45017α protein into the expression media using the yeast *Saccharomyces cerevisiae*. Two major reasons for this decision was that secreted proteins are more readily available for further manipulations (antibody production) and by passing through the secretory pathway the chances of proper folding and modifications are increased. An insight of the yeasts' secretory pathway is given in **Part I, Section 2.3**. The experimental work, cloning and expression of a full-length human P45017α in *S. cerevisiae* is described in **Part II**. Optimisation of expression conditions, determined by P45017α mRNA levels, were performed and on the basis of the results some disadvantages concerning this system have been formulated.

In **Part III**, the extracellular expression of the full-length P45017α, b_5 and aromatase in an industrial yeast, *Pichia pastoris*, was investigated. Various PCR strategies were employed to amplify the different cDNAs followed by cloning in reading frame into suitable expression vectors. The plasmids, harbouring the P45017α, b_5 and aromatase cDNAs, were subsequently integrated into the genomic DNA of two *P. pastoris* strains. Growth of different recombinant strains was monitored and expression levels of P450 enzymes were determined, using immunoblotting, in the intracellular and extracellular

Stellenbosch University http://scholar.sun.ac.za

fractions. The results of the studies are discussed and shortcomings, using *P. pastoris* as an export system for expressed P450 enzymes, are pointed out.

In order to investigate the intracellular expression of a full-length human P45017α in *P. pastoris,* cloning, integration and culturing of a suitable strain are carried out. Methods are developed and evaluated for the preparation of microsomes from *P. pastoris* expressing human P45017α. In addition, protocols are developed and evaluated for the purification of P45017α from the *P. pastoris* microsomes. These experiments are described in **Part IV**.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 2: LITERATURE REVIEWS

SECTION 2.1: An insight in structure, function and role of two key steroidogenic enzymes P45017α and aromatase

2.1.1 Cytochromes P450

The absorbance spectra of all heme proteins reveal a characteristic peak at ~ 420 nm called the Soret peak. Factors that cause this specific absorbance are the chemistry of the heme and its attachment to the protein. The conjugated bond system of heme can be influenced by changes in the nearby protein, resulting in a change of electron distribution in the heme which, in turn, will cause a shift in the Soret peak. A common method used to demonstrate this phenomenon is when carbon monoxide (CO) is bubbled through a solution of reduced P450. The CO competes with oxygen for binding to the iron and causes a peak to appear at 450 nm instead of 420 nm, hence the name "cytochrome P450". This spectral shift provides the most useful method for detecting and measuring cytochrome P450 in biological samples (4, 5).

Cytochrome P450 consists of an apoprotein of 30.000-60.000 Daltons (Da) together with a protoheme. The heme iron, which is attached to the protein by hydrophobic and Coulombic forces, can exist in either of the two-valence states — pentacoordinate or hexacoordinate. Four of the valencies are associated with the four pyrrole nitrogens of the heme. The fifth bond involves a sulphur atom in the protein taking the form of a thiolate bond with cysteine as seen in Figure 2.1.

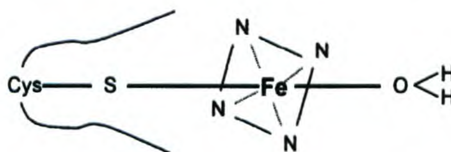

Figure 2.1: Conformational representation of heme binding in cytochromes P450.

Stellenbosch University http://scholar.sun.ac.za

It is suggested that the active site of these enzymes, which must include the heme group to which the substrate oxygen binds, take the form of hydrophobic crevices. The four bonds to the pyrrole nitrogen are lying in the same plane as the heme ring, hence planar and the fifth and sixth bonds are at a right angle to this plane, hence axial and distal. It is still uncertain as to what function the sixth ligand has.

2.1.1.1 The mechanism of cytochromes P450

The major role of the cytochrome P450 family of enzymes is to convert a lipophilic substrate to a hydrophilic substance by adding a new hydroxyl group, thereby removing the harmful substrate from the body. Such a reaction is accomplished by means of a monooxygenation, which can be represented as follows:

The P450-dependent enzymes, which are reduced by electrons from reduced pyridine nucleotides (NADPH), activate molecular oxygen. One atom of the oxygen is inserted into the substrate to form a hydroxyl group while the other is incorporated into water (6). The electron donors, involved in the reduction of cytochrome P450-dependent enzymes, differ between the endoplasmic reticulum (ER) and the mitochondria. In the ER electrons from NADPH are transferred to cytochromes P450 via a membrane-bound NADPH-P450 reductase (OR), containing both Flavin adenine dinucleotide (FAD) and Flavin mononucleotide (FMN), whereas in mitochondria this transfer is mediated by a FAD-containing reductase and a ferrodoxin-type iron-sulphur protein that is loosely bound to the inner membrane (7).

Like other enzymes the P450 enzyme releases the product so that it can bind to the next substrate. The reduced P450 enzyme then is restored to its original state (i.e. the ferric state) in order to oxygenate the next substrate. A schematic representation of the catalytic cycle of the cytochromes P450, which consists of eight reactions, is given in Figure 2.2. During the first step the substrate binds to P450 to form a substrate-enzyme complex. The binding elicits a conformational change in the P450 protein, i.e. a perturbation of heme iron spin-state equilibrium (from low spin to high spin) (8), thereby

UNIVERSITEIT STELLENBOSCH
BIBLIOTEEK

Stellenbosch University http://scholar.sun.ac.za

lowering the P450 redox potential which facilitates the next step where the first electron (donated by NADPH) reduces the iron (ferrous state) of the substrate-enzyme complex (9). In step three the reduced heme of the complex binds oxygen returning the iron to its low-spin state. A more reactive form of substrate-enzyme complex (i.e. a superoxide attached to the ferric iron) is subsequently formed in step four followed by a second electron transfer process in step five thereby reducing the ferric iron back to the ferrous state. Steps five to eight, involving the mechanism of oxygen insertion into substrates, is still under investigation and up to date three possible mechanisms of oxygen insertion to produce oxy-iron intermediates have been proposed (10, 11, 12) and are summarised by Lewis (13). The substrate is subsequently hydroxylated in step eight. The last step of the cycle is the release of the hydroxylated product from the enzyme. A reaction enhanced by the hydroxyl group that creates repulsive forces in the hydrophobic crevice of the enzyme that houses the heme group and the substrate. The free oxidised P450 is ready to start the cycle.

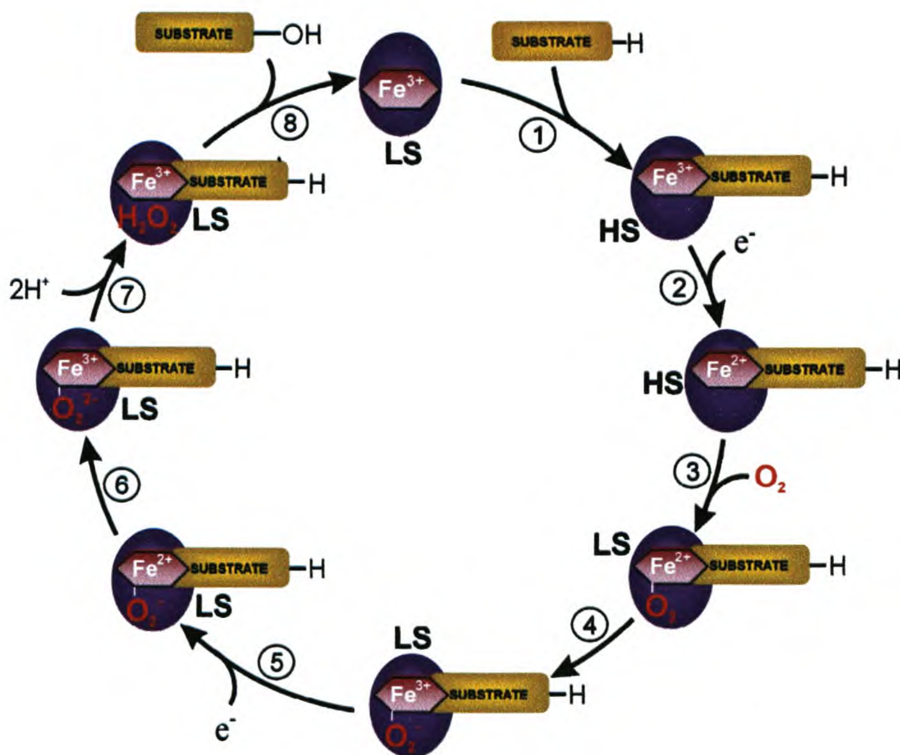

Figure 2.2 Schematic representation of the cytochrome P450 catalytic cycle showing the oxygenation of a substrate (yellow) by a cytochrome P450-dependent enzyme (purple). LS: low-spin, HS: high spin. Adapted from (13).

Stellenbosch University http://scholar.sun.ac.za

2.1.1.2 Membrane topology of cytochromes P450 together with their electron donors

Cytochrome P450-dependent monooxygenase systems of mammals are found associated with the mitochondria and ER. In order to understand the full regulation of biosynthetic reactions involving P450 in these membranes, the arrangement of the different P450 proteins, located in the membrane and their interaction with other membrane proteins, have to be understood. Up to date, we can distinguish between three different cytochrome P450-dependent monooxygenase systems, which are shown in Figure 2.3.

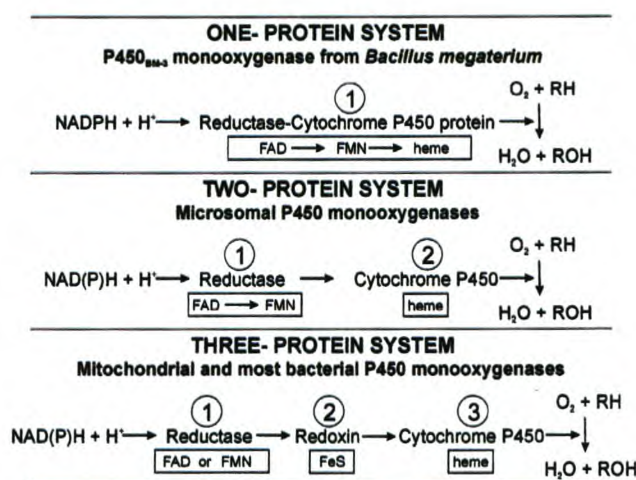

Figure 2.3 Comparison of three different P450 monooxygenase systems. The bacterial $P450_{BM-3}$ monooxygenase from *Bacillus megaterium* uses one protein to reconstitute the monooxygenase activity, whereas microsomes use two and the mitochondria three proteins. Adapted from (14).

The first system, a single protein P450-dependent monooxygenase system that was isolated from *Bacillus megaterium,* consists of a single soluble polypeptide complex containing FAD, FMN and a P450 heme moiety. The mechanism of electron transfer has been discussed in (14). The second system is the two-protein P450-dependent monooxygenase system found in the ER of higher animal cells as well as in most other eukaryotes, including yeast and is by far the largest class of cytochromes P450. Since P45017α and aromatase are part of this microsomal P450 system, it will be a point of focus further on in this section. The third monooxygenase system is the three-protein component type that occurs in the mitochondria of mammalian cells (normally those involved in steroid metabolism) and in those bacteria shown to contain P450-dependent monooxygenase systems. Here the transfer of electrons is mediated by two proteins before reduction of the P450 enzyme. One protein is a reductase (containing FAD)

Stellenbosch University http://scholar.sun.ac.za

which interacts directly with NADPH and the second protein, a small ferrodoxin-like protein (adrenodoxin). After transfer of one electron to adrenodoxin, which associates with the reductase, it dissociates from the reductase and binds to the P450 in a mechanism known as the "shuttle mechanism" (15). The mitochondrial cytochromes P450 are intrinsic membrane proteins while the associated electron-transfer components can be isolated in soluble form. On the other hand the bacterial proteins, including the P450 components, are generally soluble (16).

2.1.1.2.1 Microsomal P450 monooxygenases

2.1.1.2.1.1 Mechanism of action of cytochromes P450 and their electron carriers

The microsomal P450 system has only one electron-transfer protein that interacts with NADPH and transfers electrons directly to the various cytochrome P450 species. This protein, OR, contains both FAD and FMN. FAD receives the first electron from NADPH whereas FMN serves as the analogue to the redoxin (used in the mitochondrial three-protein system shown in Figure 2.3) by interacting with the P450 enzyme and reducing it. Numerous studies have been carried out to evaluate those factors that influence the interaction of the OR with the P450 since protein-protein interaction is essential for electron transfer [207]. The ratio between these two proteins must be 1:1 to ensure electron transfer (17). It has been proposed that the docking is affected between the positively charged amino acids on the P450 and the negatively charged amino acids on the reductase. This was confirmed by Shen and Kasper who showed unique clusters of acidic amino acids on the surface of the reductase near the FMN-binding domain, that might serve as binding site for interaction with cytochrome P450 (18).

2.1.1.2.1.2 Mechanism of membrane insertion and protein retention

Sequences of the microsomal cytochromes begin with a highly hydrophobic segment of about 20 amino acid residues that is followed by a short polycationic segment. The microsomal cytochromes P450 fall in the integral membranous proteins class that may be inserted into the ER by an amino-terminal signal peptide, but one that is not cleaved and also serves as an anchor. The signal-anchor sequence is followed by a stop-transfer signal, which is a short segment of light cationic amino acid contents that halts the

Stellenbosch University http://scholar.sun.ac.za

translocation of the rest of the protein across the membrane (19, 20). To note is also the absence of N-linked glycosylation or any Lys-Asp-Glu-Leu (KDEL)- like sequence at the carboxyl terminus, which is a characteristic of the ER retention signal of soluble ER proteins.

The mechanism of the cytochromes P450 ER translocation machinery is complex and was illustrated well by Roux *et al.* (21). There is a signal recognition particle (SRP), an SRP-receptor/docking protein, a chaperone protein and a ribosome receptor. The molecular determinants for the insertion of the signal and the mechanism of insertion have not yet been determined. There are possible mechanisms of how the mammalian cytochromes P450 are inserted into the ER, which are well reviewed (22, 23). Two of these postulations were recently shown to be correct (Figure 2.4). The P450 is located in the ER either bitopic (i.e. touching both leaflets of the membrane) or polytopic (multiple membrane crossing). The former is also called type I signal-anchor (SA-I) and the latter type II signal-anchor (SA-II). In both of these suggestions the carboxyl terminus is in the cytosol. Nelson and Strobel (24) concluded that at the amino terminus was a helical hairpin and a large, globular heme domain with the plane of the prosthetic group parallel to the membrane. Therefore the amino-terminus together with the heme domain would be situated on the cytosol side of the ER. This would comply with the polytopic postulation.

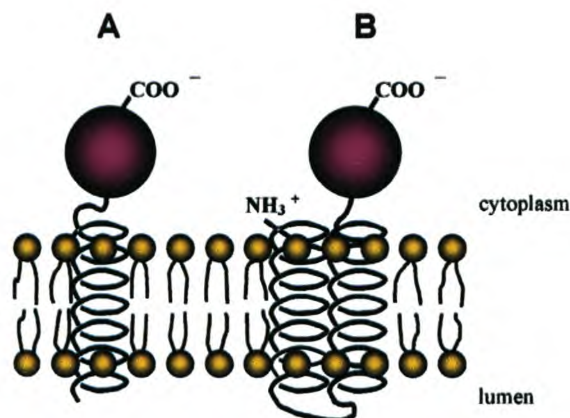

Figure 2.4 Two possibilities of the integral membrane topology for the mammalian cytochromes P450 of the ER. (A) shows the scenarios where one membrane anchor is used type I signal-anchor (SA-I) and **(B)** P450 uses the two membrane anchor inserted as a "hairpin" or SA-II.

Stellenbosch University http://scholar.sun.ac.za

The b_5 is anchored to the membrane by a carboxyl-terminal hydrophobic segment, which penetrates the membrane ER exposing its carboxyl-terminus in the lumen (25, 26).

2.1.2 Enzymes in the steroidogenic pathways and steroid hydroxylases

Steroidogenesis from precursor cholesterol involves several complex enzymatic reactions, most of which are mediated by specific cytochromes P450 located primarily in the adrenal cortex, gonads and placenta. The first step of steroidogenesis, illustrated in Figure 2.5, is the conversion of cholesterol to pregnenolone by the mitochondrial cholesterol side-chain cleavage enzyme, cytochrome P450scc, located on the matrix side of the inner mitochondrial membrane (27). This conversion is accomplished using a three-step process involving the hydroxylation of C20 and C22 followed by cleavage of the C-C bond yielding the C21 steroid pregnenolone. The availability of cholesterol for this reaction is the determining step to initiate steroidogenesis and is regulated at the level of cholesterol flow from the outer mitochondrial membrane to the inner mitochondrial membrane mediated by the steroidogenic acute regulatory protein (StAR) (28). The mechanism and site of action of StAR is still under investigation (29).

The next step in steroid biosynthesis is the metabolism of the intermediate hormone pregnenolone, which is the key substrate of androgen and oestrogen synthesis in the gonads and mineralocorticoids and glucocorticoids in the adrenal cortex. Two different enzymes, 3β-HSD (3β-hydroxysteroiddehydrogenase Δ^5-isomerase) and P45017α can metabolise pregnenolone and their competition for pregnenolone will determine the flow of steroids through the different pathways. In addition, P45017α can carry out both hydroxylase and lyase reactions, which makes this enzyme a centre player in the branch point in steroid hormone synthesis. P45017α can direct pregnenolone to mineralocorticoids (neither activity), to glucocorticoids (17α-hydroxylation but not $C_{17,20}$ cleavage) or to sex steroids (after both 17α-hydroxylation and $C_{17,20}$ cleavage), as seen in Figure 2.5.

Stellenbosch University http://scholar.sun.ac.za

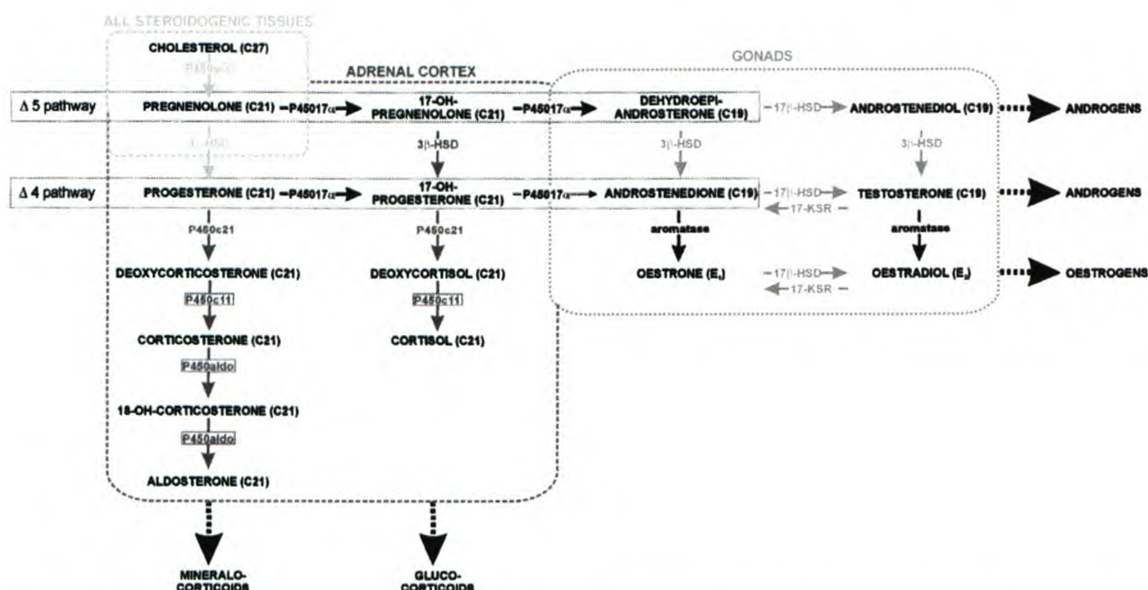

Figure 2.5 Pathway of initial steps of human steroid hormone biosynthesis in the adrenal cortex and the gonadal cells. The arrows indicate substrate conversion to product by the enzyme listed next to the arrow. The enzymes boxed in are located in the mitochondria whereas the rest are found in the ER. The physiological function of each steroid end product is indicated by the dotted arrows. Pregnenolone, the precursor for progesterone, androgen and oestrogen synthesis, is used by P45017α or 3β-HSD in a competing reaction. If pregnenolone is first used by P45017α, it is oxidised to 17-hydroxypregnenolone (17α-hydroxylase activity) and then to a Δ^5 steroid, dehydroepiandrosterone, by subsequent cleavage of the side chain (17,20-lyase activity). It is referred to as the Δ^5 pathway. If, however, pregnenolone is first used by 3β-HSD, the Δ^4 steroid, progesterone, is formed and subsequently metabolised to 17-hydroxyprogesterone (17α-hydroxylase activity) and then to the androgen precursor androstenedione by P45017α (17,20-lyase activity), hence the Δ^4 pathway. This latter 17,20-lyase activity, indicated by a thin dotted arrow, is extremely low in humans, primates, sheep and cattle and therefore androstenedione is synthesised via the Δ^5 pathway. P450scc: (cytochrome P450 side chain cleavage enzyme), P45017α: (17α-hydroxylase), P450c21: (21-hydroxylase), P450c11: (11β-hydroxylase), P450aldo: (aldosterone synthase), 3β-HSD: (3β-hydroxysteroid dehydrogenase), 17β-HSD: (17β-hydroxysteroid dehydrogenase), 17-KSR: (17-ketosteroid reductase).

The use of the pathways to produce these different steroidal hormones are determined by the different steroidogenic tissues which express specific steroidal enzymes in particular zones as summarised in Table 2.1 .

The gonads are involved primarily in production of sex hormones including oestrogens and androgens. This is achieved by the two sequential reactions catalysed by P45017α. The first reaction is the 17α-hydroxylation of pregnenolone to 17-hydroxypregnenolone and the second reaction is the 17,20-lyase activity, wherein the C_{21} steroid 17α-hydroxypregnenolone is converted to the C_{19} androgen, dehydroandrostenedione (DHEA) and subsequently androstenedione. In the testis, androstenedione is converted to testosterone and 5α-dihydrotestosterone by a series of non-P450-dependent enzymatic

Stellenbosch University http://scholar.sun.ac.za

reactions, whereas in the ovary it may be converted to oestrone catalysed by aromatase, located only in the ovary (30). A second oestrogen, oestradiol, can be synthesised either from testosterone or oestrone, mediated by aromatase and 17β-hydroxysteroid dehydrogenase (17β-HSD), respectively. The cytochrome P450-dependent 21-hydroxylase (P450c21) is not expressed in gonads (excluding many fish species), therefore intermediates are not channelled to the mineralocorticoids and glucocorticoids.

The adrenal cortex produces mineralocorticoids and glucocorticoids within a single gland, due to the broadest spectrum of steroid hydroxylases, including P450c11 and P450c21, which are solely available in adrenal tissue (Table 2.1). Progesterone, an intermediate in corticosterone and aldosterone biosynthesis and 17α-hydroxyprogesterone, an intermediate in cortisol biosynthesis, are both metabolised by P450c21 to deoxycorticosterone and deoxycortisol, respectively (Figure 2.5).

Table 2.1 Steroidogenic enzymes expressed specifically in different mammalian cell types. P450c11 and P450c21 are expressed solely in the adrenals and aromatase solely in the ovary.

	Adrenal cortex		Ovary					Testis		Placenta
Cytochrome	Zona fasciculata/ reticularis	Zona glomeru losa	Theca interna	Granulosa (preantral follicle	Luteinized theca (corpus luteum)	Luteinized granulosa (corpus luteum)	Granulosa (prevulatory follicle)	Leydig (interstitial tissue)	Sertoli	Trophoblast
P450scc	+	+	+	–	+	+	+	+		+
P45017α	+[1]	–	+	–	+	–	–	+		+[2]
3β-HSD	+	+	+	+	+	+	+	+		–
P450c11	+	–[3]	–	–	–	–	–	–		–
P450c21	+	+	–	–	–	–	–	–		–
P450aldo	–	+	–	–	–	–	–	–		–
Aromatase	–	–	–	–	–	+[4]	+	–	+	+
17β-HSD	+	?	+	–	–	–	–	+		–
Major Steroid products	**Gluco-Corticoid** Androgen	**Mineralo-corticoid**	**Androgen**	**Oestrogen**	**Progesterone** Androgen	**Progesterone** Oestrogen	**Progestins** Progesterone	**Androgen**	**Androgen**	**Oestrogen** Progesterone

[1] Not expressed in rat and rabbit adrenals, but found in hamster.
[2] Significant activity found in certain species including sheep, but human only P450scc and aromatase (30).
[3] Present in rat.
[4] Expression dependent on species.

Stellenbosch University http://scholar.sun.ac.za

The deoxycorticosterone and deoxycortisol moves from the ER to the mitochondria where the cytochrome P450 aldosterone synthase (P450aldo) and cytochrome P450 11β-hydroxylase (P450c11) catalyses the production of aldosterone and cortisol, respectively.

Apart from producing mineralocorticoids and glucocorticoids, the adrenal tissue of higher mammals is capable of extra-gonadal C19 steroid synthesis. DHEA and its sulphate conjugate, DHEA-S, are the most abundant steroids produced by the human adrenal, but their role and function has not been established other than as precursors for sex steroid biosynthesis.

2.1.2.1 Role and mechanism of P45017α in gonadal and adrenal steroidogenesis

2.1.2.1.1 Role of the 17α-hydroxylase and C17,20-lyase reaction

One of the important abilities of P45017α is that it can metabolise more than one substrate. Hence P45017α can act on both pregnenolone and progesterone (Figure 2.5). 3βHSD can act also on pregnenolone, 17-OH pregnenolone and DHEA. Therefore there are two alternative ways by which C19 steroids may be derived, which are known as the Δ^5 and Δ^4 pathways, respectively. In the Δ^5 pathway, pregnenolone is primarily utilised by P45017α followed by 3βHSD, resulting in the synthesis of 17-OH pregnenolone, DHEA via C17,20-lyase and finally androstenedione. In the Δ^4 pathway, the initial metabolism is the conversion of pregnenolone to progesterone by 3βHSD, which can be further converted by P45017α to 17-OH-progesterone and then to androstenedione. As was mentioned before, the steroid flux through each pathway is determined by 1) the competition between P45017α and 3βHSD for pregnenolone and the intermediates of Δ^4/Δ^5 metabolism and 2) the relative efficiency (substrate specificity) of the hydroxylase and lyase metabolism of the Δ^4 and Δ^5 steroids (31). In reproductive biology, however, the 17,20-lyase activity is more important since there are significant differences between species and tissue, based on the specificity of the P45017α towards its substrate. Porcine P45017α has 17,20-lyase activity for both Δ^5 (17α-hydroxypregnenolone to DHEA) and Δ^4 (17α-hydroxyprogesterone to androstenedione) pathways (32, 33), as do the rat (34), hamster (35), trout (36) and shark enzymes (37). But in guinea pig the P45017α enzyme

Stellenbosch University http://scholar.sun.ac.za

seems to be incapable of metabolising 17-OH pregnenolone to DHEA (38). In human, bovine and baboon adrenals, it appears that P45017α selectively produces DHEA in greater quantities than androstenedione (39, 40, 41), In hamster adrenals, P45017α expression favours androstene production, based on experimental results obtained with adrenal cell suspensions and microsomal preparations (42, 43, 44). However, P45017α is not expressed in the adrenal glands of rats (34) and mice (45), although in the gonads of these species P45017α has 17,20-lyase activity preferentially producing androstenedione (34). It is, however, possible for human and baboon P45017α to convert progesterone to 16α-hydroxyprogesterone (46, 41).

2.1.2.1.2 Mechanism of 17α-hydroxylase

The reaction mechanism of 17α-hydroxylation of pregnenolone, catalysed by P45017α, is well understood and documented in (47). An iron-bound peroxy intermediate is heterolytically cleaved to form activated oxygen, known as "oxene". The oxene then takes away a hydrogen atom from the substrate which results in the formation of the equivalent of an iron-bound hydroxy radical as well as a substrate-based radical

2.1.2.1.3 Mechanism and regulation of 17,20-lyase activity

The studies of Akhtar *et al.* as well as Swinney (48, 47) support the hypothesis that P45017α catalyses carbon-carbon bond cleavage resulting in the formation androstenedione from 17α-hydroxyprogesterone via a ferric-peroxy-dependent pathway. The ferric peroxy intermediate could then rearrange heterolytically, via the Baeyer-Villiger rearrangement, to form a hemiacetal or homolytically gem-diol form (47).

The 17α-hydroxylase and 17,20-lyase activities of human P45017α can be regulated independently (49) and the latter is influenced by three factors — (i) by the binding and interaction of redox the partner OR (50, 51); (ii) by serine/threonine phosphorylation of P45017α (52) and (iii) by the presence of b_5 (53). The exact role of b_5 in human P45017α is, however, controversial.

Stellenbosch University http://scholar.sun.ac.za

To date it is not clear how or if b_5 regulates human P45017α activities *in vivo.* The augmentation of the 17,20-lyase activity of P45017α by b_5 has been observed *in vitro*, but could not be seen in COS-1 cells. Similar effects can be seen in the guinea pig adrenal microsomes (54). In human adrenals the addition of b_5 promoted hydroxylase and Δ^5-lyase activities (48) On the other hand no effect of b_5 was observed in the studies on sheep adrenal P45017α (55). The conclusion of these findings is that the hydroxylase/lyase activities are species-specific and that the ratio of 17,20-lyase to 17α-hydroxylase activity of P45017α is regulated by the availability of reducing equivalents flowing to the enzyme. This can be increased by increasing the molar concentration of electron-donating redox partners, such as OR or possibly b_5.

Auchus *et al.* (56) proposed the function of b_5 which is illustrated in Figure 2.6 and concludes that the 17,20-lyase activity is not mediated by electron transfer from b_5 and that b_5 exerts an allosteric effect on the P45017α-OR complex. This suggests that b_5 interacts primarily with the P45017α-OR complex and not with P45017α alone. This was confirmed by three other independent laboratories (54, 57, 58). The structure of the interaction of P45017α with OR is still uncertain. Some clues are, however, provided from the structural interaction between the rat OR and P450-BMP.

A crevasse that extends down to the face of the P450-heme opposite the substrate-binding pocket (59), is lined on the one side with positively charged residues from the J' and K helices which participate in electrostatic pairing with negatively charged residues in OR. Geller *et al.* (50) showed with molecular modelling that the P45017α also has this similarly located crevasse of positively charged residues which interact with OR. The FMN protein of rat OR is also situated at the base of the concave cleft which is connected to the FAD domain via a flexible hinge. When the OR approaches the concave binding site of the P45017α, the FMN moiety must move 90° to ensure binding (50). Geller also believes that since the oxidative scission of the C_{17}-C_{20} bond of 17α-hydroxypregnenolone appears to require stronger constraints on the active-site topology of P45017α than 17α-hydroxylase reaction, the b_5 plays a role in the optimisation of the geometry of P45017α-OR complex (Figure 2.6). This postulation of b_5 being an allosteric effector has been strengthened by the observation that serine phosphorylation

Stellenbosch University http://scholar.sun.ac.za

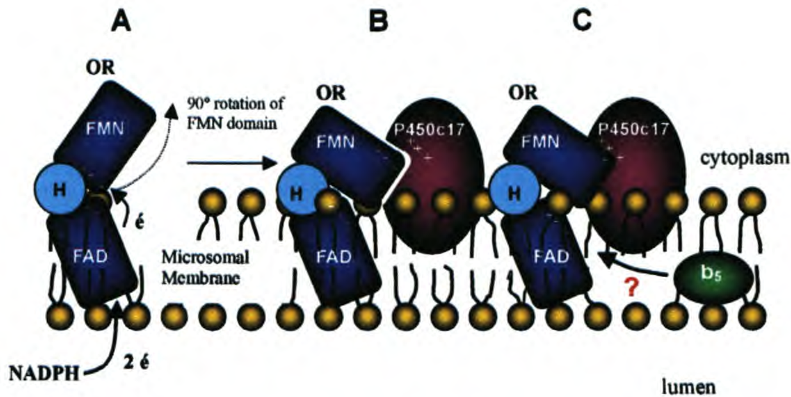

Figure 2.6 Proposed mechanism of the electron transfer and conformational arrangement of the microsomal NADPH-reductase with P45017α. (A) Two electrons from NADPH are donated to the flavin adenine dinucleotide (FAD) domain of NADPH-oxidoreductase (OR) in the microsomal membrane. The electron passes on to flavin mononucleotide (FMN). **(B)** In order for the translocation of the electrons to the P45017α the OR with its two moieties (connected together by a connecting domain or hinge region (H)) has to dock onto the P45017α binding site. This is achieved by a 90 ° rotation of the FMN domain. The P45017α-OR complex is adequate to support 17α-hydroxylation, but rarely catalyses the 17,20-lyase reaction. Therefore b_5 is needed. **(C)** The presence of b_5 facilitates the orientation of P45017α and OR in a way that is required for the more stringent 17,20-lyase reaction. This results in an electron transfer from OR to P45017α and subsequent catalysis. The precise binding site for the b_5 is still not known. This mechanism has been modified from (56).

of P45017α selectively increases 17,20-lyase activity (52) by increasing electron transfer and coupling efficiency.

2.1.2.2 Role and mechanism of aromatase in gonadal steroidogenesis

Aromatase is the cytochrome P450 responsible for the conversion of androgens to oestrogens, shown in Figure 2.5. Oestrogen synthesis is critical for normal physiological processes such as expression of secondary sexual characteristic, establishment and maintenance of pregnancy and maintaining bone mineral homeostasis (60). In humans, aromatase is expressed in steroidogenic tissues, e.g. ovary (granulosa cells), placenta and testis (Table 2.1 , Sertoli and Leydig cells) and in a number of extragonadal tissues including, breast, muscles, adipocytes, skin and in various sites in the brain and (61, 62). Recent evidence suggests that aromatase may also be capable of xenobiotic metabolism (63).

Stellenbosch University http://scholar.sun.ac.za

2.1.2.2.1 Reaction mechanism of aromatase

Aromatase converts androstenedione to oestrone and testosterone to oestradiol. The aromatase mechanism involves a unique 3-step process, which is shown in Figure 2.7. The first two oxidative steps are thought to be typical cytochrome P450 hydroxylations. A theory originally proposed by Akhtar *et al.*, suggests the third step to be a ferric peroxide attack to remove the aldehyde, followed by the aromatisation of the A ring.

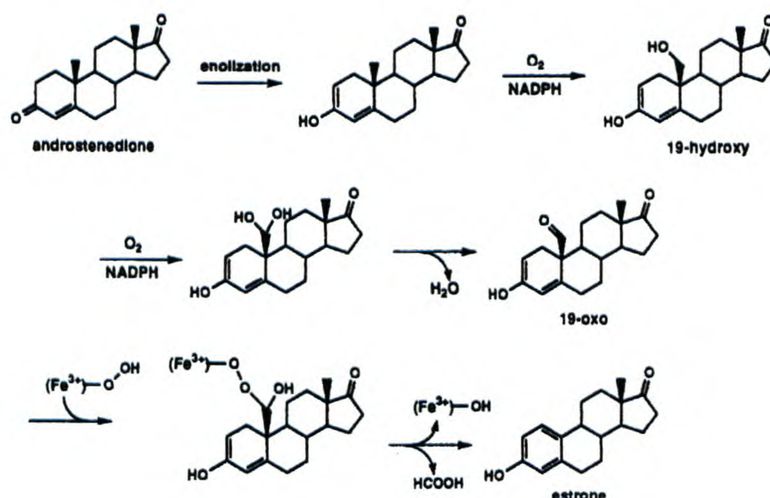

Figure 2.7 Overview of the reaction mechanism of aromatase showing the conversion of androstenedione to oestrone. Adapted from (64).

2.1.3 Sequence and structural similarities of P45017α and aromatase genes

During the past decade, cDNAs for all the major steroidogenic enzymes have been isolated and it has been shown that almost every steroidogenic enzyme is encoded by only one gene (65). The single P45017α gene (CYP17) has been characterised in humans and it contains eight exons located on human chromosome 10 (65). The human and bovine P45017α amino acid sequence consists between 508-509 amino acids in contrast with rat P45017α, which contains 507 amino acids.

Evolutionarily, the oldest of the steroid hydroxylase genes is CYP19 which encodes aromatase. The derived coding region of aromatase clearly indicates that this

Stellenbosch University http://scholar.sun.ac.za

polypeptide belongs to a member of the cytochrome P450 superfamily of genes. Generally, no greater than 30% sequence homology was found between aromatase and other forms of cytochromes P450, indicating that aromatase, in common with other steroidogenic P450 species, belongs to a separate gene family within the overall superfamily, designated cytochrome P450XIX (66). The length of 503 residues and the calculated exact molecular weight for equine aromatase is 57800 Da.

Cytochrome P45017α and aromatase have similar structural features, which are common to all cytochromes P450, sharing the most highly conserved regions in the 200 residue carboxyl terminus shown in Figure 2.8. There is a heme-binding region which has a cysteine that is believed to be the fifth coordinating ligand of the heme iron. Upstream is the second most conserved region in all P450 species, the I-helix, which has been suggested to be involved in substrate binding by forming the substrate-binding pocket proximal to the heme prosthetic group (67, 68) that contain Ozols (residues 349-371 in K-helix) and aromatic (residues 407-418) (69, 70). In addition, near the amino-terminal end a highly hydrophobic region (residues 1-20), present in all microsomal P450 species and believed to be the membrane-spanning domain, is found. However, it is the amino-terminal side of the aromatase that differs from most microsomal membrane cytochromes P450. The membrane-anchoring domain, comprising highly hydrophobic amino acid residues at the amino-terminus, as seen with P45017α, is not located at the amino-terminus in aromatase, but is situated 20 amino acids downstream from the methionine start codon, shown in Figure 2.8. Interestingly, the first 20 amino acids of the aromatase sequence have been identified as an N-linked glycosylation consensus sequence with asparagine at position 12 the site of glycosylation. Aromatase of human placenta has been shown to have attached sugar residues (71, 72, 73). It is believed that the first 20 amino acid sequence with the sugar residues is protruding into the luminal space of the ER membrane with the hydrophobic region (residue 20 to 40) acting as the membrane-spanning domain (64).

Stellenbosch University http://scholar.sun.ac.za

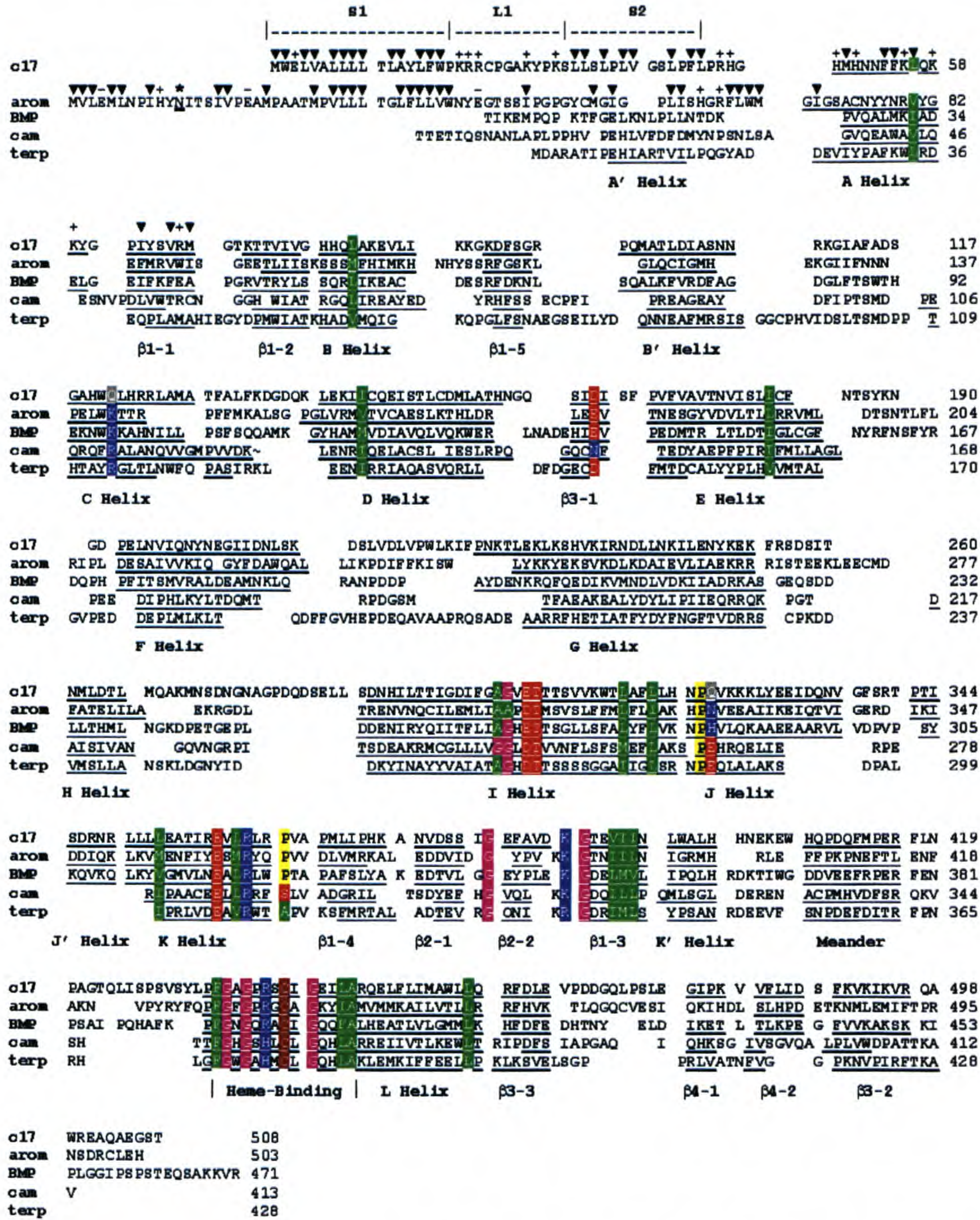

Figure 2.8 Sequence alignment of human P45017α and aromatase with the structural alignments of $P450_{BMP}$, P450cam and P450terp. Residues that comprise the α-helices (alphabetically identified) and β-sheets (numbered) are underlined and the most conserved residues (the C-helix tryptophane, catalytic threonine in the I-helix, the ExxR sequence in the K-helix and the absolutely conserved cysteine in the heme-binding region) are in colour. The amino-terminal sequence of P45017α and aromatase are shown with symbols above to indicate hydrophobic (▼), positively charged (+) and negatively charged residues (-). The putative membrane-spanning domain of P45017α is indicated by the lined sequence (S1 and S2). Aromatase has a glycosylation site at position 12.

Stellenbosch University http://scholar.sun.ac.za

This overview clearly shows that these steroidogenic cytochromes P450 (P45017α, aromatase and b_5) play an essential role in the biosynthesis of the mineralo-and glucocorticoids as well as the sex hormones. It is, however, obvious that a large amount of research, concerning these P450 enzymes, still needs to be conducted in order to obtain a precise insight into their role, localisation and function in their physiological environment. In addition, these cytochromes P450 have been implicated in various diseases, such as polycystic ovarian syndrome (74, 75), congenital adrenal hyperplasia (76) and various cancers (77), although a detailed description falls outside the scope of this thesis. In order to resolve these issues, pure P450 enzymes need to be produced via heterologous expression systems, which are discussed in the next section.

Stellenbosch University http://scholar.sun.ac.za

SECTION 2.2: Comparative study of heterologous expression systems for high-level expression of cytochromes P450 — Potential Applications of human steroidogenic cytochromes P450 expressed in *Pichia pastoris*

2.2.1 Introduction

Heterologous expression of recombinant cytochromes P450 provides unparalleled versatility for investigating different aspects of these mixed-function oxidases. One can examine, for example, the enzymatic properties of individual forms of P450 in the absence of significant background enzymatic activities (78) and reconstruct enzymatic pathways involving specific forms of P450 in cells that do not normally contain these pathways (79). One can carry out site-directed mutagenesis studies that will facilitate investigations into the catalytic role of specific amino acids (80). Furthermore, metabolic reactions of P450 enzymes can be assessed even when these enzymes are expressed at low levels in human tissue.

An important application for the heterologous expression of cytochromes P450 is to achieve large-scale production of recombinant enzymes, which is the main focus of this chapter. Once sufficient amounts of a specific form of P450 can be produced, downstream applications such as antibody production and crystallisation studies can be carried out. The crystallisation of mammalian P450 enzymes is vitally important since the basic structures of these intrinsic membrane proteins have not been resolved. The overexpression of unmodified[1] full-length P450 enzymes remains an enormous challenge in the field of biotechnology and successes have been limited. The main contributing factors to this shortcoming are the hosts and the type of expression systems, which will be discussed in this chapter.

---

[1] The modifications are in the context of truncation or alteration (e.g. nucleotide substitution or deletion) of the P450 encoding sequence at the amino terminus. Epitope tags such as FLAG and histidine are fused to the protein and do not alter the native P450 encoding sequence.

Stellenbosch University http://scholar.sun.ac.za

Although COS cells and vaccinia viruses are widely used to study recombinant P450 enzymes and their activities, they are not suitable for large-scale expression and production. The transfection efficiency of COS cells can be as low as 10%, and this system is therefore not capable of producing sufficient quantities of P450 enzymes for biophysical studies. Vaccinia viruses has been used extensively for drug metabolism studies with regard to the hepatic P450 enzymes, since they harbour a fair amount of ER, OR, b_5 and microsomal epoxide hydrolase (81). However, the expression levels are also inefficient for large-scale production of P450 enzymes.

Heterologous expression systems such as *Escherichia coli*, baculovirus and *Saccharomyces cerevisiae* which are capable of high expression levels, focus on xenobiotic-metabolising P450 enzymes, because of their substrate specificity in drug metabolism. The application of these expression systems has shifted from protein production to toxicology studies which are of vital importance in drug development (82, 83). Bacterial mutagenicity assays have been developed for screening newly developed drugs for genotoxicity and human P450 expression vectors used in *E. coli* are now being transformed into *Salmonella typhimurium*, the classical Ames tester strain (84, 85, 86). A baculoviral system has been developed to ascertain the function of various enzymes by modulating the toxic effects of chemicals in an eukaryotic environment (87). The worldwide focus has thus shifted to drug-metabolising enzymes and their potential applications in drug development.

Although some success has been achieved with regards to the functional expression of steroidogenic human cytochromes P450, the expression levels and yields are relatively low. However, high-level expression is critical in the large-scale production of these proteins for biophysical studies, such as the protein-membrane interactions, site-directed mutagenesis and crystallisation to be accomplished. *E. coli* expression systems were found to produce the highest yield of P450 enzymes followed by baculoviral cells and *S. cerevisiae* cells. Studies done by Barnes *et al.* indicated that the yield of expressed bovine P45017α between *E. coli*, baculovirus and *S. cerevisiae* was 412, 88 and 2.5 nmol/ml culture, respectively (88). However, every system has its limitations, and will be discussed in this chapter.

Stellenbosch University http://scholar.sun.ac.za

It seems that in the P450 research arena the abovementioned expression systems have reached limitations and many strategies which have the potential of increasing expression levels via co-expression and using multiple-copy clones, were unsatisfactory. However, leading the way into the new millennium, novel expression systems are becoming available and showing great success in producing high yields of various proteins (89, 90). The methylotrophic yeast, *P. pastoris*, has been employed successfully in the production of high levels of a broad range of heterologous proteins (91). Factors, which have contributed towards the success of this system, include high cell density cultures, vector stability, strong and tightly regulated promoters and the ease with which applications can be scaled up for fermentor production.

The aims of this review are to present the advantages and limitations of expression systems — conventional bacterial, baculoviral and yeast systems and non-conventional *P. pastoris* yeast systems with respect to high-level expression and yields of P450 enzymes. Factors, which may increase expression levels and affect overexpression, will be discussed. These include the host strain, specific strategies for co-expression systems, the influence of promoters on transcriptional control, formatting the initiation site of the cDNA creating a strong initiation context and removing potential secondary structures, adapting fermentation procedures with different organisms with regard to copy-number and vector stability thus increasing cell density and the use of signal peptides to direct extracellular secretion thereby facilitating purification. Finally, we will show the potential applications in the use of *P. pastoris* cells expressing full-length steroidogenic human P45017α, aromatase and b_5.

Stellenbosch University http://scholar.sun.ac.za

2.2.2 Factors influencing high P450 expression levels in heterologous expression systems

2.2.2.1 **The host**

The most striking advantages of using **bacteria** as a heterologous expression system are that bacteria are genetically well defined, grow rapidly, DNA manipulations are easy, fermentation technology is well developed and, most importantly, bacteria are a proven host for may heterologous proteins. Many results have shown that bacteria are generating large quantities of enzymatically active proteins. However, in P450 research, heterologous expression of an active P450 enzyme was not achieved unless NADPH-reductase and b_5 were present. These enzymes do not occur in bacteria and although they have an alternative electron transfer system (92, 93), it was found desirable to co-express OR and b_5 with the P450 enzyme (94, 95, 96). Numerous attempts to enhance the electron transport to P450 enzymes using various co-expression methods have been investigated and the P450 enzymes expressed to date, in bacteria have been summarised in Table 2.2 A. It clearly shows that, although the activities of expressed P450 enzymes increased, the overall yield was not improved unless the amino terminal of the P450 enzyme was modified significantly.

Protein expression in **baculovirus** can produce high levels of cytochromes P450 as shown in Table 2.2 B, although relatively labour intensive and time consuming. Using insect cells, the system is an attractive candidate for expressing large functional amounts of recombinant P450, including bovine P45017α (88). Insect cells contain both ER and mitochondria, but are deficient in all electron-transport components and hemin (97), required for the expression of catalytically active P450. Therefore these must be added to the cell extracts, making it a costly procedure. Another shortcoming of baculoviral cells is that the enzymes are produced transiently in the host cells, thereby increasing variability between different batches of preparations (98).

The yeast ***S. cerevisiae*** has several advantageous features establishing it as an important tool in the expression of foreign proteins for research, industrial or medical applications.

Stellenbosch University http://scholar.sun.ac.za

As a. GRAS (Generally Regarded As Safe) food organism, it is highly acceptable for the production of pharmaceutical proteins. *S. cerevisiae* can be grown rapidly to a high cell density on simple media resulting in the high-level production of soluble cytosolic as well as secreted proteins (99). Yeasts are free of lipopolysaccharides and associated problems of pyrogenicity, which are problems encountered with bacteria. Furthermore, *S. cerevisiae* is a useful host for the expression of mammalian genes, especially cytochromes P450, as yeast cells contain microsomal membranes, OR and b_5 resulting in membrane integration and catalytic stability. It was the first system used successfully by Oeda *et al.* (100, 101) to express a mammalian P450. Novel strains of *S. cerevisiae*, integrating the genes encoding either human OR or hamster OR have been designed and many more P450 enzymes have since been produced in yeast as listed in Table 2.2 C.

The methylotrophic yeast, ***Pichia pastoris***, has become the yeast of choice since *S. cerevisiae* has major drawbacks — the absence of strong, tightly-regulated promoters, the need for a fed-batch fermentor to reach high cell densities and the occurrence of hyperglycosylation (99). The *P. pastoris* heterologous expression system has become one of the most popular systems used in the production of recombinant proteins, including enzymes, proteases, protease inhibitors, receptors, single-chain antibodies, and regulatory proteins. Several general reviews of this system have been published giving a broad insight into *P. pastoris* (102, 103, 104, 105). The application of *P. pastoris* in the P450 research field has gained popularity, as reports show that this expression host is well suited in producing unmodified P450 enzymes at moderate to high yields, as listed in Table 2.2 D. GS115 (*his4*) and the wild-type X-33 *P. pastoris* strains, which have a higher growth rate using methanol as the sole carbon source, are the most commonly used *P. pastoris* strains (106). KM71 (*his4*, *aox1::arg4*), is a slower growing strain utilising methanol at a lower rate, has been shown to be advantageous if overexpressed proteins are detrimental to the heterologous host (107). In addition, protease deficient strains such as SMD1168 are also available (105).

2.2.2.2 Co-expression of cytochromes P450 with OR

In **bacteria** co-expression of P450 enzymes and OR as a fusion protein has become a common approach to express active P450 enzymes and was first introduced by Fisher *et*

Stellenbosch University https://scholar.sun.ac.za

al. when expressing bovine P45017α or rat P450 4A1 fused to rat OR (400-700 nmol/L)(108). However, expressing P450 enzymes as a fusion protein with rat OR in the *E. coli* showed little or no metabolic activity (109, 7, 110), although moderate expression levels (50-250 nmol/L) were obtained, as shown in Table 2.2 A. In an attempt to increase catalytic activity and expression levels of a human P450 enzymes in bacteria, a bicistronic system was developed by Dong *et al.* (111), which overcame the problem of the insufficient electron transport from NADPH to heterologously expressed P450. A bicistronic vector was designed by cloning a bicistronic construct downstream of a single promoter (*tac*). This construct consisted of a microsomal P450 cDNA encoded by the first cistron (first ribosome binding site) and the auxiliary protein OR cDNA encoded by the second cistron (second ribosome binding site). The vector was used for independent co-expression of P450 and NADPH-P450 reductase (bicistronic construct) and was shown by Shet *et al.* to produce levels up to 200 nmol/L culture when co-expressing the steroidogenic bovine P45017α and rat OR. (112). However, for most P450 enzymes co-expressed with the bicistronic vector, an increase in activity was achieved, but the yields of the recombinant P450 proteins decreased significantly, seen in Table 2.2 A.

There is strong evidence that to increase the turnover of some substrates the bicistronic vector is unsatisfactory and the addition of b_5 essential. Voice *et al.* co-expressed P450 3A4, OR and b_5, in an *E. coli* system (113) and showed that the turnover of the substrates increased in the presence of b_5 in the whole cell and bacterial membrane fraction. Although the stability of a spectrally active enzyme was also prolonged, the expression levels of P450 3A4 was reduced by ~ 50%.

Another attempt to increase expression levels in bacteria is to use double promoter constructs as described by Blake *et al.* (114). The modified human P450 3A4 cDNA and human OR cDNA, each independently under control of the *tac* promoters, produced P450 3A4 expression levels of approximately 200 nmol/L culture. This level was lower than when P450 3A4 was expressed alone or as a fusion protein with OR which yielded 370 and 230 nmol/L culture, respectively (115, 116). Many microsomal human liver P450 enzymes have since been expressed in *E. coli* and in *S. typhimurium* and are listed

Stellenbosch University http://scholar.sun.ac.za

in Table 2.2 A. However, the data show that the expression levels of P450 without OR are higher that P450 expression with double promoter constructs.

A final approach to increase expression levels of P450 enzymes in bacteria was to introduce two different plasmids carrying respective human P450 cDNA and human OR cDNA into *E. coli* (JM109) cells simultaneously, first shown by Pritchard *et al.* (117). Expression levels of P450 2D6 in *E. coli* whole cells determined by CO-difference spectra were 381 nmol/L culture.

Baculovirus expression offers high-level expression (300-1000 pmol/mg total cell lysate) of both P450 and OR, using a single transfer vector. However, in a co-expression system, using a dual expression plasmid (i.e. P450 cDNA and OR cDNA located in one plasmid controlled by two independent promoters), the catalytic activity of expressed P450 enzymes can be increased although the yields of spectral P450 decreased around 10 fold when compared to the expression of P450 cDNA without OR cDNA (118). Wang *et al.* have shown that although an imbalance of the molar ratio between the human P450 2E1 and the OR of 1:5 occurs when using a dual expression virus, 170 pmol/mg membrane protein was produced (119). Another approach to increase expression levels was to co-infect Sf9 cells with two baculoviruses, one containing human P450 2E1 cDNA and the other containing human OR cDNA, which yielded expression levels of human P450 2E1 at 400-800 pmol/mg cell proteins. (119).

S. cerevisiae cells contain endogenous OR and b_5 activity, although low levels and availability can impact negatively on the specific activity of expressed P450 enzymes (120, 121). Co-expression of OR with P450 enhances the P450-dependent monooxygenase activity in yeast as was shown by Murakami *et al.* when expressing rat P450 with OR as a fusion protein in *S. cerevisiae* (122). Although the focus was not on increasing the yield of recombinant P450, the expression of bovine P45017α, rat P450 1A1 and bovine P450c21 fused to yeast OR (123, 124, 125, 126), led not only to the gaining of knowledge in applications of the fusion expression systems but also to invaluable insight into the interaction between cytochrome P450 and OR (1-200 pmol/mg microsomal protein).

Stellenbosch University http://scholar.sun.ac.za

Alternative strategies for co-expressing OR with P450 enzymes in *S. cerevisiae* is to either use a single plasmid containing multiple expression cassettes, i.e. each recombinant cDNA has its own promoter (127, 128, 129), or transforming yeast cells with multiple vectors (130). However, these yeast expression systems are mitotically unstable, due to the large plasmids burdening the cells resulting in low cell growth. Pompon *et al.* therefore developed a yeast expression system to optimise the redox environment, since earlier yeast expression systems showed interference from endogenous OR and b_5 (131). The gene encoding endogenous OR located on the yeast chromosomal DNA, under the control of the GAL10-CYC1 promoter, was disrupted by site-directed integration of the human OR gene. In another strain the OR gene was substituted with the expression cassette containing human b_5. Co-expression of the associated electron-transfer proteins by integration of galactose-inducible expression cassettes into the yeast genome under the control of an efficient and externally controllable promoter, enhanced specific P450 3A4 activity 73-fold in comparison with the activity observed in a wild-type strain (132, 133). However, the production level of P450 1A1 in the microsomal fraction did not improve and the level of P450 3A4 decreased (134, 135). Ohgiya *et al.* followed the same principal of integration of the expression unit into the yeast genome, inserting hamster OR at multiple-copy sites into the yeast retrotransposon *Ty*, resulting in a 20-times increase of P450 reductase activity compared to the parent yeast (136).

To date co-expression of OR with P450 enzymes in ***P. pastoris*** has not been investigated. Preliminary results from our research group (not published) show high P45017α hydroxylase and lyase activity in whole cells, indicating that the endogenous yeast OR and b_5 electron transport to the P45017α enzyme is adequate.

In summary, different strategies employed in the co-expression of the electron transport equivalents, have lead to enhanced efficiency of electron transfer from OR to P450 enzymes, increasing the catalytic activity of the expressed enzymes albeit at lower yields.

2.2.2.3 Co-expression of P450 enzymes with chaperones

The heterologous expression of P450 enzymes in *E. coli* is enhanced by ethanol, which triggers the heat shock response (associated with protein folding through molecular chaperones) and some C-group antibiotics, which trigger the cold shock response (translational regulation by protein synthesis inhibitors (137, 138). Inoue *et al.* reported elevated expression levels of P450 3A7 in the presence of molecular chaperone GroEL, known to assist the correct folding of proteins in *E. coli* (139). This chaperone, co-expressed with a modified human aromatase gene, also led to stable expression levels of 400 nmol/L culture in *E. coli* (140). The use of another heat shock protein, HSP40, enhanced expression levels and catalytic activity of P450 1A2 in bicistronic format (141). However, although the co-expression with chaperones increased the production of several P450 enzymes, the success of this strategy appears to be protein specific (142).

2.2.2.4 Transcriptional control: vectors and promoters

Heterologous expression levels in any organism are dependent on the type of vector and copy-number used to introduce foreign DNA in the host. Expression levels are influenced by the type (induced or constitutive) and strength of the promotor as well as gene stability. In the last decade various promoters have been investigated and their application in the heterologous expression of cytochrome P450 will be discussed.

One of the first promoters used in ***E. coli*** systems to express of large amounts of bovine P45017α was the powerful T7 phage promoter (143). However, most of the recombinant proteins were found in inclusion bodies since the expression rate of the recombinant proteins exceeded the capability of the cell to synthesise these ancillary elements thus failing to fold correctly and accumulate irreversibly in inclusion bodies (IB). Many proteins produced in *E. coli* accumulate in IB as a denatured form or can be present as dimers and higher-molecular-mass multimers. Tedious isolation protocols, including denaturation techniques, need to be applied compromising high yields of correctly folded protein. To date, the most preferred promoter for cytochrome P450 expression is the *tac-tac* promoter of the pCWori+ vector (Table 2.2 A.

Stellenbosch University https://scholar.sun.ac.za

In **baculoviral** systems a conventional strong polyhedrin promoter is commonly used in the expression of heterologous proteins in Sf9 cells. However, high levels of expression of recombinant P450, coupled to a strong promoter, can be detrimental to correct folding and heme incorporation resulting in an inactive and insoluble apoprotein. To overcome these shortcomings a moderate or weaker promoter was employed by Paine *et al.* (118).

Initially the alcohol dehydrogenase promoter (*ADH1*) located in the pAAH5 vector (101, 144, 145, 146) was the promotor of choice in ***S. cerevisiae*** systems but many alternative promoters have since been used. The acid phosphatase (*PHO5*) (147, 148) and metallothionine (*CUP1*) promoter in the inducible vector Yep13 (149, 150,151) were able to express a number of different P450 enzymes at low concentrations in *S. cerevisiae* expression systems, as summarised in Table 2.2 C. The use of a medium strong constitutive phosphoglycerate kinase (PGK) and glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter produced moderate levels of P450 enzymes (40-100 pmol/mg microsomal proteins) in the yeast strains AH22 (152, 153, 154, 155, 156) and W303 (56). Urban *et al.* improved the yield of expressed P450 enzymes in yeast by developing the GAL10-CYC1 hybrid promoter which remains one of the most powerful, tightly-regulated, inducible promoters of *S. cerevisiae* (132, 131). This promoter is induced more than 1000-fold by the addition of galactose but strongly repressed by glucose. In cultures grown in glucose medium maximal induction can only be achieved following depletion of glucose at the end of the exponential phase and the addition of galactose during the stationary phase (99). In yeast, tightly regulated promoters therefore separate the growth phase from the induction phase minimising stress and avoiding the selection of variants expressed at low levels often seen during large-scale culture fermentations.

The improved expression levels of P450 enzymes cannot be attributed to the GAL10-CYC1 hybrid promoter alone since the plasmid copy-number in the yeast cells also plays a vital role. Pompon *et al.* used *E. coli*-yeast shuttle vectors based on the 2μ autonomously replicating vectors, pYeDP (which are stably maintained at 50 to 200 copies per cell (133, 157). The successful production of human cytochrome P450 1A1 was achieved in *S. cerevisiae* YME2 by incorporating the cDNA encoding the enzyme into the yeast 2μ episomal plasmid pDP34 (158, 159). Unfortunately yeast cells, which

Stellenbosch University http://scholar.sun.ac.za

harbour episomal vectors need to be grown under selective pressure to ensure high levels of expression and stability. The cells therefore have to be grown on a selective medium (selective defined), resulting often in retarded growth, which in turn decreases the yield of expressed protein considerably.

Integrating vectors (YIp) contain yeast chromosomal DNA which permits the expression units, selectable marker and bacterial replicon to be stably integrated via a single crossover integration resulting in the duplication of the chromosomal target sequence. Yeast cells can therefore be grown for sixty generations without selective pressure in a rich medium (160). The copy number remains constant in *S. cerevisiae* and independent of culture conditions, it is very low (1 copy per cell) and expression levels therefore depend only upon promoter efficiency.

The compromise between high, stable integration and low copy number can be circumvented in the ***P. pastoris*** multi-copy expression system. Cleavage of the *P. pastoris* vector within a sequence, shared by the host genome, stimulates homologous recombination events, which target integration of the vector to either the *his4* or the *aox1* loci efficiently. The latter is shown in Figure 2. 2.9. Alternatively, single-copy integration may also be achieved replacing the *AOX1* gene (transplacement) via double homologous recombination, depicted in Figure 2.10. However, during these single homologous recombinations, multi-copy integration can arise (1-10% frequency) and multi-copy transformants containing over 20 tandem repeats can be generated (161). In most cases, multi-copy integrants are more likely to express higher levels of the recombinant protein in comparison to the single-copy strains (162).

P450 proteins can be expressed in *P. pastoris* controlled by the strong and highly regulated alcohol oxidase (*AOX1*) promoter, which can maintain P450 genes in a suppressed expression mode on a non-methanolic carbon source minimising selection for non-expressing mutant strains during cell growth. The promoter is switched on by shifting to methanol feeds. The AOX1 promoter therefore tightly regulates expression at transcriptional level. RNA is undetectable in cells cultured on carbon sources such as glucose, glycerol or ethanol, but in methanol-grown cultures the AOX1 message represents ~ 5% of total polyA$^+$ RNA producing 30% of the total soluble protein (163).

The *AOX1* promoter produced moderate levels of a plant P450 (128 pmol/mg microsomal proteins) in yeast strain X-33 (164) and higher levels of the human P45017α (300 pmol/mg microsomal proteins) in yeast strain GS115 (165).

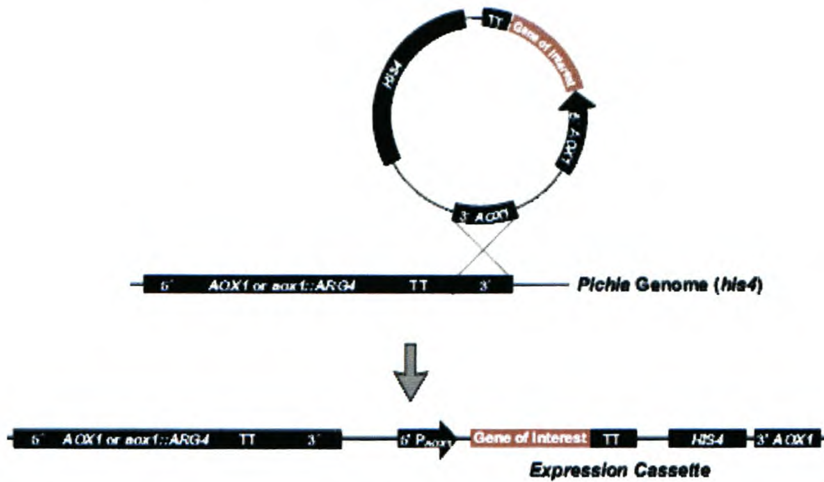

Figure 2. 2.9 Schematic representation of a gene insert event at *AOX1*. In GS115 cells, gene insertion events at the *aox1* locus arise from a single crossover event between the *aox1* locus in the chromosome and the *AOX1* gene on the vector. One or more copies of the vector are inserted at the *aox1* locus. By linearising the recombinant vector at a restriction enzyme site located in *AOX1* gene, Mut+ recombinants can be conveniently generated depending on the host strain used. (From Ref. 166)

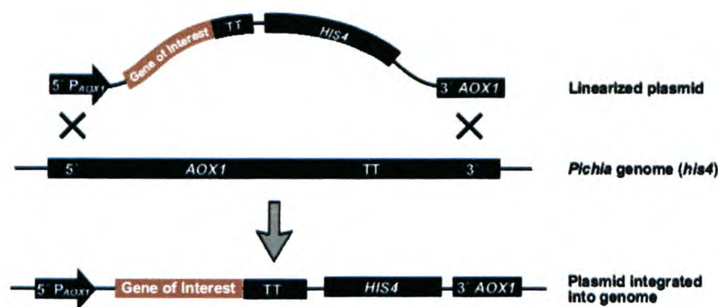

Figure 2.10 Schematic representation of a gene replacement event at *AOX1*. In a his4 strain such as GS115, a gene replacement (omega insertion) event arises from a double crossover event between the *AOX1* promoter and 3′ *AOX1* regions of the vector and genome. This results in the complete removal of the *AOX1* coding region (i.e. gene replacement) and the gain of an expression cassette containing P_{AOX1}, the gene of interest, and *HIS4*. (From Ref. 166)

2.2.2.5 Translational efficiency: strong initiation sites, secondary structures and codon usage

Factors, such as codon usage, GC content, secondary structure formation, potential splice sites and favourable translation initiation codon context in a given mRNA, can influence the efficiency of translation, often resulting in low levels of heterologously expressed proteins. The parameters of these factors differ between host organisms and studies, conducted in COS (167), *S. cerevisiae* (168, 169) and *E. coli* cells (170), have shown that certain codons and consensus sequences around the initiation site are recognised most efficiently by the most abundantly available tRNA species (171, 172, 173). Based on this knowledge, P450 researchers have modified the amino terminal of the P450 cDNA which, in many cases, lead to increased levels of heterologous protein expression. These strategies are summarised below.

The expression levels of full-length P450 enzymes are very low in **bacterial** expression systems and efficient expression was achieved with modifications of the lipophilic amino-terminal sequence. In *E. coli*, parameters such as codon usage preferable to bacteria and the minimisation of secondary structure of mRNA ensure good translational initiation context (174, 175). Since the expression of the native form of bovine CYP17 cDNA failed in E coli, Barnes *et al.* increased the expression levels by modifying the 5'-end, illustrated in Figure 2.11 (176). The second codon was changed from TGG (Trp) to GCT (Ala) to ensure a preferred second codon for expression of the *LacZ* gene (177). Furthermore, modifying codons 4 and 5 to TTA created an AT rich sequence which is a more favourable initiation sequence, since most mRNA expressed in *E. coli* has been shown to be adenosine and uridine nucleotides (178). In addition, the last nucleotides of codons 6 and 7 were changed to A and T respectively, which ensured a decrease in the synthesis of the secondary structures around the ribosome binding-site which resulting an inhibition of translation (174, 175).

When the modified bovine amino-terminal sequence (MALLLAVF) replaced the amino-terminus of human P45017α (Figure 2.11 B modified), expression of human P45017α increased by 40% (179). Numerous cytochromes P450 have since been expressed in *E. coli* by generating chimeric enzymes utilising the modified bovine sequence.

Stellenbosch University http://scholar.sun.ac.za

A

native bovine CYP17:	ATG	TGG	CTG	CTC	CTG	GCT	GTC	TTT	CTG...
	Met	Trp	Leu	Leu	Leu	Ala	Val	Phe	Leu...
modified :	ATG	GCT	CTG	TTA	TTA	GCA	GTT	TTT	CTG...
	Met	Ala	Leu	Leu	Leu	Ala	Val	Phe	Leu...

B

native human CYP17:	ATG	TGG	GAG	CTC	GTG	GCT	CTC	TTG	CTG...
	Met	Trp	Glu	Leu	Val	Ala	Val	Phe	Leu...
modified :	ATG	GCT	CTG	TTA	TTA	GCA	GTT	TTT	CTG...
	Met	Ala	Leu	Leu	Leu	Ala	Val	Phe	Leu...

Figure 2.11 Modifications of the amino-terminus for the efficient expression of bovine P45017α (A) and human P45017α (B) in *E. coli*. The red indicates the nucleotide changes introduced via PCR mutagenesis.

Other studies have attempted to reduce the number of amino acid changes, replacing only the second codon with alanine, and found different levels of expression, suggesting that additional factors influencing translational initiation are required for high-level expression in *E. coli*. (180, 181, 182). However, Shimada *et al.* observed no consistent correlation between expression levels of human P450 1B1 with the potential for secondary structure formation in the mRNA transcripts or with the presence of Ala/GCT in the second codon position (183).

Another strategy to increase expression levels in conjunction with the solubility of P450 enzymes was to truncate the highly hydrophobic amino-terminal region of the P450 (184). The first report from Li *et al.* demonstrated that deleting the amino-terminal hydrophobic sequence of amino acids 2-24 (P450c7Δ2-24) increased the expression level by 10% and, in addition, the protein was mainly cytosolic (185). Sagara *et al.* expressed a truncated bovine P45017α (Δ2-17) in *E. coli*, and although not able to integrate into the membrane, remained functional in *E. coli* (186). It was therefore concluded that the hydrophobic amino terminus was not essential for the folding pathway. In contrast, in mammalian cells the folding pathway leading to functional P45017α appears to require a signal anchor sequence since the same truncated P45017α is found to be inactive (187). Larson *et al.* found that the removal of amino acid residues 3-29 did not prevent a functional P450 2E1 binding to the membrane (188). It seems, that the latest trend in attempting to overexpress cytochromes P450 successfully in bacteria is to delete the hydrophobic amino-terminal sequence and replacing it with either the 5'-MALLLAVF sequence or various favourable translation initiation codon

Stellenbosch University http://scholar.sun.ac.za

contexts such as 5'-MARQVHSSWN and 5'-MAKKTSSK. Additional altered modifications are summarised in Table 2.2 A.

Unlike *E. coli* based expression of P450 enzymes, it is unnecessary to modify mammalian cDNAs extensively for their heterologous expression in **baculoviral systems.**

Few reports dealing with amino-terminal modifications of P450 cDNA in order to increase the expression levels of P450 enzymes in ***S. cerevisiae*** have been published. A more recent study by Batard *et al.* investigated the adjustment of the amino-terminal of the coding sequence of a monocot P450 gene (P450 73A17) to a more favourable codon usage characteristic of the expression host *S. cerevisiae* (189). The data suggests that certain rare codons, including repetitive CTC triplets nearer the translation initiation site, is more critical than unfavourable codons distal to the ATG start site. Recoding the first 111 base pairs yielded P450 73A17 expression levels of 300 pmol P450/mg proteins.

However, upstream of the ATG initiation site, certain consensus sequences have shown to increase initiation and promote efficient translation in ***S. cerevisiae*** (190, 191). Sakaki *et al.* employed a short (AT)-rich sequence upstream of the ATG codon, increasing the yield of bovine P45017α twofold in *S. cerevisiae* (123). Hayes *et al.* inserted two nucleotide fragments upstream of the initiation site (192) — AACA found in several abundantly expressed yeast genes (193), while GATC was previously shown to increase the expression of P450 1A2 in yeast (194).

The criteria ensuring translational efficiency in *S. cerevisiae* also apply in ***P. pastoris***. The mRNA secondary structure around the AUG start codon may be adjusted so that AUG is relatively free of secondary structure as predicted by the RNA-fold analysis computer program.

2.2.2.6 Fermentation conditions of high cell density growth in fermentor cultures

In recent years, medium optimisation and high cell density culture techniques have been employed in an endeavour to enhance productivity of proteins in fermentation processes

Stellenbosch University http://scholar.sun.ac.za

utilising recombinant organisms. Scale-up procedures of heterologous expression systems, from a shake-flask environment to fermentors, are a complex process. Not only are fermentors ideal for the efficient production of expressed proteins, but they can be used as chemostats maintaining constant cell cultures growth, important for enzyme kinetics. Controlled conditions will contribute to reproducibility which many P450 expression systems lack. The expression of recombinant P450 enzymes has been optimised primarily for *in vivo* activity studies and the plasmids and promoters are therefore unsuitable for high cell density fermentations.

Protein production in ***E. coli*** can be increased significantly through the use of high-cell-density culture systems, but most expression plasmids used to date are unstable in long fermentation periods. In addition, expression of P450 enzymes in *E. coli* can be increased with the supplementation of cofactors, such as a heme precursor, δ-aminolevulinic acid (δ-ALA) and flavins (195, (196, 197, 198,199). This would, however, lead to an increase in the costs of large-scale fermentations.

Baculovirus expression system also requires the addition of hemin to the culture medium for the production of functional P450 enzymes. The baculoviral system can be scaled up to larger volumes for the use in fermenter systems. However, since the consequence of infection is cell death, it is not possible for relevant protein production to be maintained in continuous cultures.

In ***S. cerevisiae***, heme is the prosthetic group of different hemoproteins or enzymes involved in important physiological functions of yeast cells, such as respiration (cytochromes of the mitochondrial respiratory chain), detoxification (peroxidase and catalases), sterol and unsaturated fatty acid metabolism (b_5 and cytochromes P450) and sporulation. It was shown that yeast expressing bovine P45017α is capable of acquiring heme, even under heme-limiting conditions. However, competition for heme between hemoproteins incorporating heme and the cells' own vital ergosterol biosynthesis, was reported (200). The hierarchy of fulfilment of heme requirements may however be avoided by the supplementation of the medium with hemin.

Stellenbosch University https://scholar.sun.ac.za

Scaling-up into production fermentor systems (110-liter), utilising *S. cerevisiae* strains D12 and AH22 transformed with the vector pAAH5 harbouring P450 3A4 or P450 2C10 cDNA, was described by Guengerich *et al.* (201). Results indicated that the transformants could only be maintained at low cell density culture due to plasmid loss, producing low P450 2C9 and P450 3A4 levels of at 2.25 nmol/L and 2.0 nmol/L cell cultures, respectively. Nishihara *et al.* subsequently showed high-level production of the bovine P45017α protein in *S. cerevisiae* (GRF18) under the control of the GAL10 promoter (202). In a 30-L fermentor P45017α, determined by CO-difference spectra on whole cell assay, was increased 4.7-fold (37 mg/L) compared to small-scale shake-flask cultures.

The industrial yeast ***P. pastoris***, together with the *AOX1* promoter, is a unique system in large-volume high-density fermentor cultures. The promoter is efficiently transcribed in cells exposed to methanol as sole carbon source but is strongly repressed under most other growth conditions. This aspect is important since the selection for non-expressing mutant strains is minimised during the growth phase, ensuring high cell densities without induction. Unlike most regulated promoters described in *S. cerevisiae* systems, the *AOX1* promoter is regulated or "switched on" by methanol alone and remains "switched off" even under limiting growth conditions (e.g. carbon starvation) which normally results in the derepression of other promoters relying on the carbon levels. *P. pastoris* is therefore grown on glycerol containing media, a repressing carbon source which will not, even at very low levels of glycerol, lead to expression of the *AOX1* promoter. Once the glycerol has been depleted, methanol must be added to ensure induction of *AOX1* expression, allowing the cells in a fermentor environment to grow in a high cell density prior to the addition of the inducer. Large-volume continuous cultures therefore can be maintained at cell densities in excess of 100 grams/litre. Our research group is currently investigating the optimisation of heterologously expressed human P450 enzymes in fermentors, monitoring various fed-batch and continuous culture schemes (203, 204, 205).

Stellenbosch University https://scholar.sun.ac.za

2.2.2.7 **Fusion of cytochromes P450 to signal peptides and secretion**

Expressing proteins in the periplasmic space or culture media via a signal sequence can ensure simple, cost-effective downstream purification processes which ultimately increases protein production. The oxidising environment of the periplasm and secretory organelles facilitates proper protein folding and, in addition, the signal peptide is removed and degraded during translocation of the polypeptide across the inner membrane, leaving the authentic amino-terminus of the target protein intact (206).

This strategy was first used by Prichard *et al.*, whereby high levels of unmodified P450 enzymes were expressed in ***E. coli***, by fusing translational amino-terminals to bacterial leader sequence, *pelB* and *ompA* (207). The *omp*-3A4 fusion resulted in higher levels of expression and improved recoveries of active P450 enzyme in bacterial membrane fractions in comparison to yields obtained with the modified amino-terminal bovine P45017α cDNA.

The ***P. pastoris*** expression system is becoming more popular due to its ability to secrete heterologous proteins (208). In an attempt to increase human P450 production we have expressed various steroidogenic P450 enzymes extracellularly in *P. pastoris*. Preliminary results indicate low expression levels in the medium necessitating further investigation (165).

Stellenbosch University http://scholar.sun.ac.za

2.2.3 Genetically-engineered human steroidogenic cytochromes P450 from P. pastoris - potential applications

P. pastoris, is an industrial yeast which has been optimised for large-scale production of heterologous proteins in fermentors (209). Preliminary investigations into the expression of human steroidogenic P450 enzymes in *P. pastoris* in shake-flask experiments suggest that these enzymes can be overexpressed and scaled-up into fermentors, leading to major applications indicated in Figure 2.12. The *P. pastoris* expression system can be employed to produce high yields of human steroidogenic P450 proteins which may, after successful purification and downstream processes, aid in biophysical studies. The purified proteins can also be used in processes such as raising poly- and monoclonal antibodies. Whole *P. pastoris* cells can be harvested, producing a cell line mimicking metabolic profiles specific to individual tissue types permitting *in vivo* investigations. These two products, e.g. proteins and whole cells, can generate further potential applications, which will be discussed in this section. These applications can be divided into two categories — Primary fundamental applications are defined as applications wherein the expressed protein or the expression host are utilised for basic cytochrome P450 research. Many primary applications that arise from the basic cytochrome P450 research are important to various commercial industries, especially the pharmaceutical industry. Secondary applications are defined as the transfer of knowledge obtained during ongoing P450 research at academic level to the industrial level for commercial exploitation of these enzymes.

2.2.3.1 Potential applications using pure proteins

2.2.3.1.1 Analysing structures

Steroidogenic P450 proteins, produced from genetically engineered *P. pastoris* are suitable for crystallographic structural experiments. X-ray crystallography is one of the most powerful tools available for determining the structure of proteins, since the x-ray crystal structure represents the three-dimensional conformation of a protein in a solid state.

Stellenbosch University http://scholar.sun.ac.za

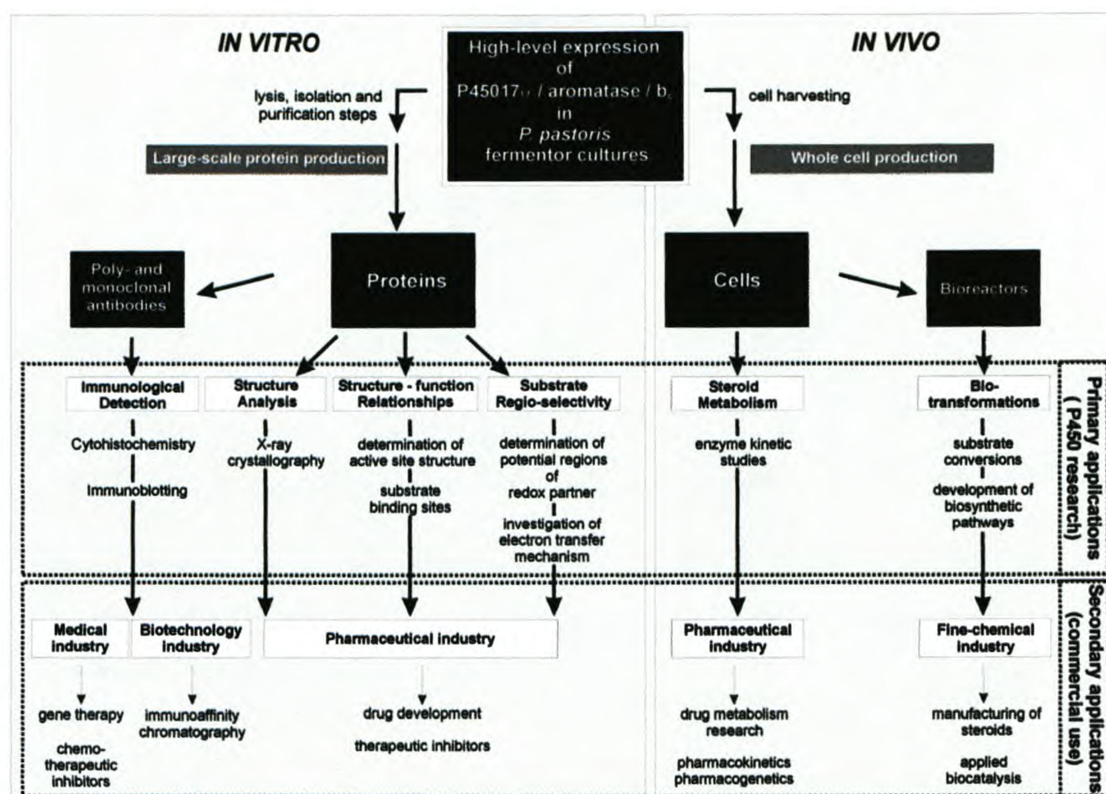

Figure 2.12 Diagrammatic representation of potential applications of human steroidogenic cytochromes P450 expressed in *P. pastoris*.

The first successful x-ray crystallography was achieved with a soluble prokaryotic $P450_{cam}$ (P450 101) (210), and subsequently with various bacterial P450 enzymes, viz. $P450_{BM3}$ (P450 102) (211, 212), $P450_{terp}$ (P450 108) (213), $P450_{eryF}$ (P450 107A1) (214), $P450_{sca2}$ (215), P450 119 (216) and $P450_{cin}$ (P450 176) (217), including a fungal $P450_{nor}$ (P450 55A1) (218). Until recently, eukaryotic cytochromes P450 could not be crystallised due to two factors. Firstly, their hydrophobic amino-terminal domain prevented the P450 proteins from being overexpressed in bacteria resulting in low yields failing to meet the requirement for crystallisation. Secondly, the microsomal P450 proteins need to be in a lipophilic environment to prevent aggregation. However, the use of nonionic surfactants such as Emulgens interfere with the crystallisation processes. Various attempts to produce a more soluble form of P450 have been made. Some research groups have removed the hydrophobic sequences (219, 198, 220) while others expressed P450 enzymes with a cleavable membrane segment removing it with specific proteases (221). Sueyoshi *et al.* replaced the amino terminus of a mouse P450 2A4 with an amphipathic 24 residue sequence (222).

Stellenbosch University http://scholar.sun.ac.za

The tertiary structures of four mammalian cytochromes P450 have been solved by X-ray analysis using these and new novel strategies to produce soluble proteins. These include rabbit P450 2C5 (223, 224), rabbit P450 2B4 (225, 226), human P450 2C8 (227) and human P450 2C9 (228). The P450 2C8 cDNA was modified by replacing the sequence encoding the first 27 amino acids with one encoding a short hydrophilic, positively charged amino-terminus (MAKKTSSKG) identical to that employed for the expression and crystallisation of P450 2C5 (220, 229). This modification of the amino-terminus of P450 2C8 converts the integral membrane protein to one that binds peripherally to membranes and that can be separated from the membrane by elevating the ionic strength of the medium. A similar modification and mutagenesis strategy was employed for human P450 2C9 protein expression (228). However, no human steroidogenic P450 protein has, to date, been crystallised. The *P. pastoris* expression system potentially allows the production of large quantities of pure P45017α proteins for x-ray crystallography.

2.2.3.1.2 Determination of substrate specificity

Human steroidogenic P450 proteins isolated from recombinant *P. pastoris* cells will contribute towards the determination of the structure/function relationships. The predicted structure of the active site, which plays a role in the **substrate specificity** of P450 enzymes, is as yet not fully understood and may be resolved from reconstituted systems in conjunction with site-directed mutagenesis studies. Important residues proximal to the active site can be altered, expressed in *P. pastoris* and activity of the mutated protein compared to that of the wild-type.

2.2.3.1.3 Elucidating electron transfer of redox partner

There is intense interest in the involvement of b_5 in the regulation of NADPH-supported cytochrome P450 monooxygenase reaction (230), specifically regarding the catalytic activity of human P45017α. Although it was shown that b_5 donates the second electron to the oxyferrous form of P450 during catalysis (231), Auchus *et al.* subsequently showed, using microsomal yeast preparations, that both holo- and apo-b_5 acts principally as an allosteric effector interacting primarily with the P45017α-OR complex stimulating

Stellenbosch University http://scholar.sun.ac.za

the 17,20-lyase, but not the 17α-hydroxylase activity (56). Similar effects have been shown for P450 3A4 (232, 57) and P450 4A7 (233), suggesting that the electron transfer to the P45017α protein is not via b_5. However, recent publication by Guryev *et al.* challenges the allosteric effector hypothesis (234). Purified human P45017α together with human b_5 expressed in *P. pastoris* cells may contribute towards the characterisation of the binding site and elucidate the mechanism of action as the precise site of action of b_5 remains unknown.

Pharmaceutical drug design will be enhanced once the three-dimensional structure of the steroidogenic P450 enzymes has been determined. A detailed knowledge of the active site will permit the synthesis of specific compounds which will bind specifically to the active site of these P450 enzymes, inhibiting their activity (235, 236, 237). P45017α and aromatase have been identified as crucial pharmacological targets for the treatment of prostatic androgen dependent diseases (benign hypertrophy and carcinoma) (238, 239) and breast cancer (240, 241), respectively. Although potent therapeutic inhibitors targeting P45017α and aromatase have been suggested (242, 243, 244) and physiochemical requirements for their inhibitory activity have been shown (245, 246), their exact interaction with the appropriate P450 enzymes are, however, predictive mathematical algorithms based on theoretical three-dimensional models.

2.2.3.1.4 Production of antibodies

Polyclonal and monoclonal antibodies (mAbs) can be raised with the successful production of human P45017α and aromatase proteins expressed by *P. pastoris*. Monoclonal antibodies, which are highly specific and often a limitless source, can be used in an array of applications for P450 research and has been reviewed by Gelboin *et al.* (247, 248). MAbs play an important role in the field of immunohistochemistry and have made possible not only the identification and localisation of specific steroidogenic proteins in selected tissue but also the quantification of these proteins. mAbs can be used in immunological detection assays such as ELISA, radioimmunoassay (RIA) and Western blotting. Furthermore, mAbs are directed to conformational or sequential epitopes on the surface of the P450 protein. Based on this principle, it was found that P450 reactions can be inhibited by mAbs binding to the epitope-specific P450 protein,

Stellenbosch University http://scholar.sun.ac.za

thus determining the location of the substrate binding site (regio-and stereospecificity) as well as the protein content.

Evidence to date indicates that oestrogens, especially 17β-oestradiol, functions as a mitogenic factor in human endometrial, ovarian and breast cancers and could potentially regulate cellular proliferation and development of these neoplasms (77). Human aromatase catalyses the final step of oestrogen biosynthesis (249) and the ideal anti-cancer strategy would be to block the terminal step in oestradiol biosynthesis without inhibiting production of other important steroids. Therefore, the application of mAbs, raised against aromatase, as a therapeutic inhibitor in oestrogen dependent cancers is an attractive alternative to synthetic chemical inhibitors (250) due to the high specificity with which mAbs recognise and bind their target antigen.

Immunoaffinity chromatography is one of the most powerful fractionation steps available in protein purification and mAbs has an application in this field of biotechnology (251). A highly selective purification and concentration method is ensured by applying a protein mixture to a specific antibody immobilised on a resin or membranous support, washing off nonspecific contaminating proteins and eluting the absorbed antigenic protein with appropriate elution agents. The use of mAbs against human aromatase and their use in purification procedures have been previously described (252, 253). Immunoaffinity chromatography has been optimised for various proteins and is another proteomics tool (254).

2.2.3.2 Potential applications using whole cells

2.2.3.2.1 Regulation of steroid metabolism

In vivo investigations using bacteria, baculoviral and yeast models represent rapid and simple methods for assessing the role of cytochromes P450 in drug metabolism under physiologically relevant conditions (255). Although whole cell assays are very useful for predictive studies, they may not reflect the true cellular complex environment where drug disposition, transport, adsorption and secretion are factors to be considered. In addition, external factors such as cell densities, pH and growth media can differ considerably from one experiment to another resulting in discrepancies. However,

Stellenbosch University http://scholar.sun.ac.za

recombinant *P. pastoris* cells can be grow under highly controlled constant conditions in an chemostat, thereby ensuring reproducibility from batch to batch. Furthermore, the characterisation of the steroidogenic enzymes with respect to kinetic parameters and optimal reaction conditions can result in superior yields by using enzyme variants that exhibit optimal activity under thermodynamically favourable conditions (256). *P. pastoris* may therefore, become a model system for steroidogenic and drug metabolism research.

2.2.3.2.2 Testing mutagenicity and carcinogenicity

The *P. pastoris* model system may be used to test the mutagenicity of chemicals once the system has been evaluated and standardised based on the enzyme kinetic levels.

2.2.3.2.3 Manufacturing of metabolite in bioreactors

At the level of the organism, knowledge pertaining to sequence and structure may be combined with metabolic pathway analysis allowing the engineering of *P. pastoris* strains that direct their carbon flow towards specific compounds. This principal of biotransformation is very useful in fine-chemical and pharmaceutical industries and certain steroids and enzymes can be manufactured at lower costs instead of using tedious and costly chemical synthesis processes (257). Hydrocortisone, an anti-inflammatory used in the treatment for rheumatoid arthritis, was recently successfully biosynthesised in recombinant *S. cerevisiae* grown on a simple carbon source (258). Based on this principal, genetically engineered *P. pastoris* whole cells, immobilised in a fixed-bed reactor or grown in stirred fermentors, would be able to produce significant steroidogenic metabolites from the complete bioconversion of substrate. Therefore, clinically marketable steroids such as oestrogens synthesised from androgens in recombinant *P. pastoris* cells, is likely possibility in the near future.

2.2.4 Conclusion

The choice of heterologous expression system for cytochrome P450 production depends on the type of investigation. If one wants, for example, to determine the enzyme activity

Stellenbosch University http://scholar.sun.ac.za

of a heterologously expressed P450 in whole cells, then the use of a yeast system would be the most likely candidate since the endogenous or co-expressed electron transfer machinery is available under most conditions to maintain activity. On the otherhand, if one wants to determine the enzyme activity in isolated microsomal environment or in a reconstituted system, baculoviral and bacterial systems are the preferred systems of choice.

We are however, interested in a suitable expression system that can a) overexpress a full-length P450 protein, b) correctly fold the translated protein, c) produce high yields and d) grow in high cell density fermentations. The most common organisms, *E. coli*, baculovirus, *S. cerevisiae* and *P. pastoris* were investigated to see which one could fulfil all these four criteria. Few reports have compared the first three expression systems based on P45017α (88) and P450 3A4 (259) expression levels and yields, but these are not truly comparable, since the cDNA of the expressed proteins were modified and optimised for each particular organism.

Although *E. coli* produces the highest levels of P450 enzymes, the cDNA of eukaryotic P450 enzymes have to be modified extensively before they can be expressed successfully. In addition, the protein folding pathway in *E. coli* differs from that in mammalian cells, leading to improper conformation, often leading to the accumulation of an inactive protein form in IB (260). Therefore, bacteria are not the ideal organisms for expressing a full-length P450 protein at high yields.

The baculovirus expression system offers high expression levels of an active P450 protein, therefore meeting three of the four criteria. However, engineering cDNA-carrying viruses is very laborious and time-consuming. Once high-level expressing cells have been identified, P450 expression levels are very difficult to maintain since the cells are transiently infected. The baculoviral system's shortcoming therefore lies in high cell density fermentations as small changes in growth conditions from batch to batch increase variability thus hampering reproducibility (98).

S. cerevisiae expression systems are preferred in the P450 field of research, since the ancillary enzymes needed for the correct functioning of the P450 enzymes are present. In

Stellenbosch University http://scholar.sun.ac.za

some cases however, the activity of the yeast OR and the availability of internal membranes in which they are anchored, limits the expression of human cytochromes P450 in yeast (261). Earlier studies using yeast expression systems were able to produce low to moderate levels of expression. However, with stable integration of plasmid DNA, several milligrams of active protein could be readily expressed and purified.

While much work has been done to modify and improve *S. cerevisiae* as an expression system, several laboratories have investigated alternative yeast hosts for the high-level production of active proteins, resulting in the identification and development of a non-*Saccharomyces* yeast. *P. pastoris* is a powerful tool for the production of foreign proteins since its *AOX1* promoter is strongly regulated and well-developed methods for classical- and molecular-genetic manipulation of the organism exist. Furthermore, the technology *for P. pastoris* growth in large high-density fermentor is well understood and the up-scaling process for commercial use can be readily achieved. Therefore the flexibility of the *P. pastoris* expression system makes it an ideal tool for P450 research in primarily fundamental applications as well as for secondary industrial applications.

Stellenbosch University http://scholar.sun.ac.za

Table 2.2 A

Summary of heterologous expression of cytochromes P450 in bacteria.

CYP[1]	Modification[2]	Co-expression method[3]	Vector	Promoter	Strain[4]	Expression level[5]		Ref.
Human								
1A1	Replace 2nd codon with Ala Max. AT content[6]	-	pCWori+	*tac*	DH5α	25	nmol/L	262
	Replace 2nd codon with Ala	f	pCWori+	*tac*	DH5α	150	nmol/L	263
	Replace 2nd codon with Ala	b	pCWori+	*tac*	DH5α	27 16	nmol/L nmol/L♦	116
	Replace 2nd codon with Ala	2 ×	pCWori+	*tac*	DH5α	310	nmol/L	264
	N/A	2 ×	pCWori+	*tac*	YG7108[7]	N/A	-	265
	N-term. MALLLAVF	2 ×	pCWori+	*tac*	TA1538[7]	54	nmol/L	266
	N term. truncation	-	pET[8], pGEX[9]	*lac,* *tac*	BL21, JM101	IB[10]	-	260
	N-term. MALLLAVF[11]	-[12]	pCWori+	*tac*	DH5α	520	nmol/L	141
1A2	N-term. Δ 1-20 aa[13], N-term. MALLLAVFL	-	pCWori+	*tac*	DH5α	700	nmol/L	267
	N-term. MALLLAVFL	-	pCWori+	*tac*	DH5α	245	nmol/L	268
	N/A	-	pCWori+	*tac*	MX100	N/A	-	269
	N-term. MALLLAVF	f	pCWori+	*tac*	DH5α	150-250	nmol/L	110
	N-term. Δ 3-13 aa, N-term. MALLLAVF	b	pCWori+	*tac*	DH5α	350 190	nmol/L nmol/L♦	116
	N-term. Δ 3-13 aa, N-term. MALLLAVF	2 ×	pCWori+	*tac*	DH5α	520	nmol/L	264
	N/A	2 ×	pCWori+	*tac*	TA1538[7]	N/A	-	270
	N/A	2 ×	pCWori+	*tac*	YG7108	N/A	-	265
1B1	N-term. Δ 2-4 aa,	-	pCWori+	*tac*	DH5α	800	nmol/L	197
	N-term. Δ 2-4 aa, N-term. MALSPNDPWPL	b	pCWori+	*tac*	DH5α	200	nmol/L	183
	N/A	2 ×	pCWori+	*tac*	YG7108	N/A	-	265
2A6	N-term. MALLLAVF	-	pCWori+	*tac*	JM109	455	nmol/L	207
	OmpA-[14], C-term. 6 × histidine tag	-	pCWori+	*tac*	JM109	193	nmol/L	207

[1] Cytochrome P450 gene nomenclature.

[2] Genetic modifications of the P450 encoding sequence at amino terminus (N-term) or carboxyl terminus (C-term.)

[3] Expression of P450 enzymes without OR (-) or with OR in certain constructs. f = fusion, b = bicistronic, t = tricistronic, 2 × = 2 genes under the control of two independent promoters, 2 p = two plasmids.

[4] All *E. coli* strains, except where indicated.

[5] nmol/L = expression level of nmol P450 enzyme per liter of culture; pmol/mg = expression level of P450 content per mg of cell protein.

[6] Maximising the AT content at the amino terminus.

♦ The yield of OR.

[7] *S. typhimurium.*

[8] Six histidine tag at carboxy-terminal.

[9] GST (glutathione demethylase) tag at the carboxyl terminus.

[10] The expressed proteins were in inclusion bodies.

[11] Replacement of the hydrophobic sequence at the amino-terminal with the modified bovine P45017α coding sequence (MALLLAVFL).

[12] Co-expression with chaperone Hsp40.

[13] deletion (Δ) of 20 amino acids (aa) at amino-terminus.

[14] Bacterial leader signal *(OmpA orPelB)* fused to the amino terminal of the full-length P450 coding sequence.

Stellenbosch University https://scholar.sun.ac.za

	N-term. Δ 3-24 aa, Replace 2nd codon with Ala, N-term. MARQVHSSWNL	2 ×	pCWori+	*tac*	DH5α	870	nmol/L	264
	N/A	2 ×	pCWori+	*tac*	YG7108	N/A	-	265
	Replace 2nd codon with Ala	-	pCWori+	*tac*	DH5α	210	nmol/L	271
2B6	N-term. MALLLAVF, C-term. 6 × histidine tag	-	pCWori+	*tac*	MV1304	25-80	nmol/L	272
	N-term. MALLLAVF	2 ×	pCWori+	*tac*	DH5α	N/A	-	273
2C2	Full-length	-	Pc2A	-	CJ236	17	nmol/L	274
2C8	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	1500	nmol/L	196
	N-term. Δ 3-27 aa, N-term. MAKKTSSKG, C-term. 4 × histidine tag	-	pCWori+	*tac*	DH5α	N/A	-	227
	N-term. Δ 3-20 aa, Replace 2nd codon with Ala, N-term. MARQVHSSWNL	2 ×	pCWori+	*tac*	DH5α	1620	nmol/L	264
	N/A	2 ×	pCWori+	*tac*	YG7108	N/A	-	265
2C9	N-term. MALLLAVF	-	pCWori+	*tac*	DH5α	5-11	nmol/L	275
	N-term. Δ 3-20 aa, Replace 2nd codon with Ala N-term. MARQSSGR	-	pCWori+	*tac*	DH5α	9-19	nmol/L	275
	Partial sequence	-	pET, pGEX	*lac, tac*	BL21, JM101	IB[10]	-	260
	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	500	nmol/L	196
	N-term. Δ 3-29 aa, N-term. MAKKTSSKGR, C-term. 4 × histidine tag							228
	N-term. Δ 3-19 aa, Replace 2nd codon with Ala	b	pCWori+	*tac*	DH5α	170 170	nmol/L nmol/L•	116
	N-term. Δ 3-20 aa, Replace 2nd codon with Ala N-term. MARQSSGRGK	2 ×	pCWori+	*tac*	DH5α	1210	nmol/L	264
	N/A	2 ×	pCWori+	*tac*	**YG7108**	N/A	-	265
2C18	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	450	nmol/L	196
2C19	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	700	nmol/L	196
	N-term. MALLLAVF	2 ×	pCWori+	*tac*	DH5α	750	nmol/L	264
	N/A	2 ×	pCWori+	*tac*	**YG7108**	N/A	-	265
2D6	N-term. MALLLAVF	-	pCWori+	*tac*	DH5α	600	nmol/L	198
	N-term. Δ 25 aa, replaced, with 6 × histidine tag	-	pDS	*T5-lac*	JM109	14.7	nmol/L	199
	N-term. Δ 18 aa, aligned, with P450 2E1	-	pCWori+	*tac*	DH5α	90	nmol/L	198
	OmpA, N-term. MALLLAVF	f	YPD2b	*tac*	DH5α	60-100	nmol/L	276
	N-term. Δ 3-25 aa, Replace 2nd codon with Ala, N-term. MARQVHSSWNL	b	pCWori+	*tac*	DH5α	130 34	nmol/L nmol/L•	116
	N-term. Δ 3-24 aa, Replace 2nd codon with Ala, N-term. MARQVHSSWN	2 ×	pCWori+	*tac*	DH5α	580	nmol/L	264
	N/A	2 ×	pCWori+	*tac*	**YG7108**	N/A	-	265
	N-term. MALLLAVF	2 p	PCWori+, pACYC184	*tac*	JM109	381	nmol/L	117
	OmpA,	2 p	PCWorI+, pACYC184	*tac*	JM109	365	nmol/L	117
2E1	Replace 2nd codon with Ala, Silent mutations 3-7 aa[6]	-	pKK233-1	*Trp-lac*	DH5α	N/A	-	277
	N-term. Δ 3-21 aa, Replace 2nd codon with Ala	-	pCWori+	*tac*	DH5α	40	nmol/L	278
	Partial sequence	-	PET; pGEX	*Lac; tac*	BL21 JM101	IB[10]	-	260
	N-term. MALLLAVF	-	pCWori+	*tac*	JM109	68	nmol/L	207
	OmpA, C-term. 6 × histidine tag	-	pCWori+	*tac*	JM109	175	nmol/L	207
	Replace 2nd codon with Ala	b	PJL2	*tac*	XL-1 blue	0.8 4.7	nmol/L nmol/L•	111
	N-term. Δ 3-21 aa, Replace 2nd codon with Ala	b	pCWori+	*tac*	DH5α	160 59	nmol/L nmol/L•	116
	N-term. Δ 3-24 aa, Replace 2nd codon with Ala, N-term. MARQVSSWNCP	2 ×	pCWori+	*tac*	DH5α	1250	nmol/L	264

Stellenbosch University http://scholar.sun.ac.za

CYP	Modification		Plasmid	Promoter	Strain	Yield	Unit	Ref.
	N/A	2 ×	pCWori+	*tac*	YG7108	N/A	-	265
3A4	N-term. Δ 3-12 aa, N-term. MALLLAVF	-	pCWori+	*tac*	DH5α	370	nmol/L	115
	Partial sequence	-	PET, pGEX	*lac, tac*	BL21, JM101	N/A	IB[10]	260
	OmpA -[14], C-term. 6 × histidine tag	-	pCWori+	*tac*	JM109	504	nmol/L	207
	PelB -[14], C-term. 6 × histidine tag	-	pCWori+	*tac*	JM109	143	nmol/L	207
	N-term. Δ 1-20 aa, N-term. MALLLAVFL	f	pCWori+	*tac*	DH5α	200	nmol/L	109
	N-term. Δ 3-12 aa, N-term. MALLLAVF	b	pCWori+	*tac*	DH5α	230 200	nmol/L nmol/L*	116
	N-term. Δ 3-12 aa, N-term. MALLLAVF	t	PB216	*tac*	JM109	82	nmol/L	113
	N-term. Δ 3-12 aa, N-term. MALLLAVF	2 ×	pCWori+	*tac*	JM109	200	nmol/L	114
	N-term. Δ 3-24 aa, N-term. MALLLAVF	2 ×	pCWori+	*tac*	DH5α	410	nmol/L	264
	N/A	2 ×	pCWori+	*tac*	YG7108	N/A	-	265
3A5	N-term. Δ 18 aa N-term. MALLLAVF	-	pCWori+	*tac*	DH5α	260	nmol/L	279
	N/A	2 ×	pCWori+	*tac*	YG7108	N/A	-	265
3A7	N-term. MALLLAVF	b	pCWori+	*tac*	DH5α	50	nmol/L	139
	N-term. MALLLAVF	-[15]	pCWori+	*tac*	DH5α	78	nmol/L	139
3A43	N-term. MALLLAVF, C-term. 6 × histidine tag	-	PSE380	N/A	TOPP3	4-28	nmol/L	280
17	N-term. MALLLAVF	-	pCWori+	*tac*	DH5α	400	nmol/L	179
	N-term. 2 aa changes	-	pCWori+	*tac*	DH5α	40	nmol/L	179
19	N-term. Δ 1-38 aa N and C-term. 6 × histidine tag	-	pET3b	*T7φ10*	BL21	N/A	-	281
	N-term. Δ 3-37 aa, N-term. MARQSFGRGKL	-[15]	pCWori+	*tac*	DH5α	240	nmol/L	140
	N-term. Δ 3-37 aa, N-term. MARQSFGRGKL, C-term 4 × histidine tag	-[15]	pCWori+	*tac*	DH5α	350-400	nmol/L	138
21			pET	*lac*	BL21			282
46A1	Replace 2nd codon with Ala, Silent mutations 3-7[6]	-	pCWori+	*tac*	DH5α	160	nmol/L	283
	N-term. Δ 3-27 aa C-term. 4 × histidine tag	-	pCWori+	*tac*	DH5α	390-650	nmol/L	284
CYB5	N-term. 4 × histidine tag	-	PT7-7	*T7*	BL21	1300	nmol/L	285
P450scc	Replace 2nd codon with Ala, Silent mutations 3-7[6]	-	PTrc99A	*N/A*	JM109	10-35	nmol/L	286
Rat								
2B1	N-term. MALLLAVF	-	pKK233-1	*Trp-lac*	XL-1 blue	35	nmol/L	287
2C11	N-term. Δ 3-7 aa, Replace 2nd codon with Ala	-	pCWori+	*tac*	DH5α	500-1500	nmol/L	288
4A1	N-term. MALLLAVF	f	pCWori+	*tac*	DH5α	400-500	nmol/L	108
4F1, 4F4, 4F6	N-term. Δ 1-23 aa, N-term. MALLLAVF, C-term. 6 × histidine tag	-	pCWori+	*tac*	DH5α	N/A	-	289
7A1	Replace 2nd codon with Ala		pKK233-1	*Trp-lac*	XL1-blue	0.2	%[16]	185
	N-term. Δ 2-24 aa		pKK233-1	*Trp-lac*	XL1-blue	2.0	%	185
CYB5	Full-length	f	pCWori+	*tac*	JM109	3000-6000	nmol/L	290
	N-term. MALLLAVF	f						**291**

[15] Co-expression with chaperone GroES/GroEL

[16] This value is probably based on apoprotein determination and does not reflect the level of expression of holoenzyme.

Stellenbosch University http://scholar.sun.ac.za

Mouse								
1B1	N-term. Δ 2-17 aa, Replace with amphipathic peptide	-	pCWori+	*tac*	DH5α	N/A	-	222
2A4	N-term. Δ 1-20 aa, N-term. MALLLAVF, C-term. 6 × histidine tag	-	pET23b	T7	BL219D3	25-60	nmol/L	292
Guinea Pig								
17	C-term. 4 × histidine tag	-	pCWori+	*tac*	JM109, DH5α, BL21	300	nmol/L	293
Bovine								
17	N-term. Δ 2-17 aa	-	pCWori+	*tac*	XL1-blue	600	nmol/L	186
	N-term. MALLLAVF	-	pCWori+	*tac*	JM1-0	412	nmol/L	176
	N-term. MALLLAVF, C-term. 4 × histidine tag	-	pCWori+	*tac*	JM109	450	nmol/L	294
	N-term. MALLLAVF, C-term. 6 × histidine tag	f	pCWori+	*tac*	DH5α	600-700	nmol/L	7
	N-term. MALLLAVF	f	pCWori+	*tac*	GM48	700	nmol/L	108
	N-term. MALLLAVF	b	pCWori+	*tac*	DH5α	150-200	nmol/L	112
Rabbit								
2B4	N-term. Δ 2-27 aa, Replace 2nd codon with Ala	-	pGEX[9]	*tac*	MV1304	300-400	nmol/L	184
	N-term. MALLLAV	-	pKK233-1	*Trp-lac*	Topp3	100	nmol/L	295
	N-term. Δ 3-21 aa, N-term. MAKKTSSKK, C-term. 4 × histidine tag	-	pKK233-3	*Trp-lac*	Topp3	800-1000	nmol/L	225
2B5	N-term. MALLLAVLL	-	pKK233-1	*Trp-lac*	Topp3	100	nmol/L	295
	N-term. MALLLAVF	-	pKK233-1	*Trp-lac*	XL-1 blue	7	nmol/L	287
2C1	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	180	nmol/L	196
2C2	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	300	nmol/L	196
2C3	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	400	nmol/L	196
	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	500	nmol/L	196
	N-term. MALLLAVF, Replace 10, 14,15[nd] codon with Gly, Leu and Leu	-	pCWori+	*tac*	XL-1 blue	400	nmol/L	296
	N-term. Δ 3-20, 4 × His C-term	-	pCWori+	*tac*	XL-1 blue	800-1200	nmol/mg	220
2C4	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	300	nmol/L	196
2C5	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	500	nmol/L	196
	NH_2 Δ 3-20, N-term. MALLLAVF, 4 × His C-term	-	pCWori+	*tac*	XL-1 blue	N/A	-	220
2C16	N-term. MALLLAVF	-	pCWori+	*tac*	XL-1 blue	350	nmol/L	196
2E1	None	-	pKKHC	*trc*	MV1304	0.3	%[17]	188
	N-term. Δ 3-29	-	pKKHC	*trc*	MV1304	0.1	%[17]	188
Canine								
2B11	N-term. MALLLAVF	-	pKK233-1	*Trp-lac*	XL-1 blue	100	nmol/L	287
Shark								
17A	N-term. MALLLAVF, C-term. 6 × histidine tag	-	pET21	*T7-lac*	BL21	-	-	37
Fish								
1A	N-term. Δ 3-31 aa, max. AT content C-term. 4 × histidine tag	-	pCWori+	*tac*	JM109	200	nmol/L	297
Plant								
71D13	N-term. MALLLAVF	-	pCWori+	*tac*	JM109	265	nmol/L	298

[17] The figure is the percentage expressed P450 of the total *E.coli* protein, measured via immunoassay and does not reflect the level of expression of holoenzyme.

Stellenbosch University http://scholar.sun.ac.za

71D15	N-term. MALLLAVF	-	pCWori+	*tac*	JM109	500	nmol/L	298
71D18	N-term. MALLLAVF	-	pCWori+	*tac*	JM109	350	nmol/L	298
73	N-term. MALLLAVF	f	PQE-6	-	DS410	N/A	-	299
79	NH_2 Δ 1-25 N-term. MALLLAVF	-	pSP19g10L	*lac*	JM109	500	nmol/L	300
House fly								
6A1	N-term. 2 aa changes	-	pSE380	*trc*	PI-13c	340	nmol/L	182
Bacteria								
101	-	-	PGK300	*lac*	JM83	-	-	301, 210
176A	None	-	pCWori+	*tac*	DH5α	2000	nmol/L	302, 217
BioI	None	-	pET 11a	*T7*	BL21, pLysS	-	-	303

Stellenbosch University http://scholar.sun.ac.za

Table 2.2 B

Summary of heterologous expression of cytochromes P450 in baculovirus.

CYP	Modification	Co-expression method[1]	Vector	Cells/ virus	Expression level[2]	Ref.
Human						
1A1						**304**
2C8						**305**
2C9	Full-length	-		Sf9	N/A	306
2C11						**307**
2D6	Full-length	-	PVL1392	Sf9	50-200 pmol/mg	118
	Full-length	2 ×	P_2Bac	Sf9	10-20 pmol/mg	118
2E1	Full-length	-		Sf9	300 pmol/mg	306
	Full-length	-	PVL1393	Sf9	1500 pmol/mg	308
	Full-length	-	BacPak	TN5/	500 pmol/mg	309
	Full-length	2 ×	P_2Bac	Sf9/	170 pmol/mg[3]	119
	Full-length	i			400-800 pmol/mg[3]	119
2J2	Full-length	-	pBlueBacIII	Sf9	100-150 nmol/L	310
3A4					460 pmol/mg	**311**
3A4	Full-length	2 ×	PAcUW51	T.ni	107 pmol/mg	312
3A7			pAcYM1			313
4A11						**314**
19	Full-length	-	pBluebac 1	Sf21/	N/A	315
	Full-length	2 ×[4]	PlucGRbac1	Sf9/ AcMNPV	N/A	316
	Full-length	-	pBluebac 1	TN5/ AcMNPV	105 pmol/mg	317
	NH_2 Δ 1-41, C-term. 6 × histidine tag	-	PVL1392	Sf9	N/A	318
	Full-length	-	pBluebac 1	TN5/ AcMNPV	500 nmol/L 950 pmol/mg	319
CYB5	Full-length	-	PVL1393	Sf9	500 pmol/mg[5]	308
Bovine						
17	Full-length	-	PVL1393	Sf9	88 nmol/L[6]	88
Rat						
2A1	Full-length	-	PAC373	Sf9/ AcMNPV	44 pmol/mg	97
	Full-length	-		Sf9	300 pmol/mg	306
2C11						**307**
2E1	Full-length		BacPak		300 pmol/mg	309
2J4			PVL1392		40-60 nmol/L	320

[1] Expression of P450s without OR (-) or with OR in certain constructs. 2 × = Dual expression system. The P450 and OR cDNAs in a single plasmid are controlled by two independent promoters., i = co-infection of two baculoviruses, one with the P450 cDNA and the other with OR cDNA.

[2] pmol/mg = expression level of P450 content per mg of total cellular protein. nmol/L = expression level of nmol P450 enzyme per liter of culture

[3] pmol/mg of membrane protein

[4] Aromatase co-expressed with luciferase

[5] Specific content of P450 per mg protein in crude cell lysate.

[6] P450 content in membranes

Stellenbosch University http://scholar.sun.ac.za

Table 2.2 C

Summary of heterologous expression of cytochromes P450 in *S. cerevisiae*.

CYP	Modification	Co-expression method[1]	Vector	Promoter	Strain	Expression level[2]		Ref.
Human								
1A1	Full-length	2 ×	pHE10	*GAPDH*	YHE2	3	mg/ml	127, 128
	Full-length	i	pYe DP	*GAL10-CYC1*	W303.1B	43	pmol/mg	132, 134
	Full-length	-	pHE10/11	*GAPFL, POH5*	YHE2	156	pmol/mg	159
	Full-length	-	pCK-1	*PGK*	AH22	43	pmol/mg	152
	Full-length	2 ×	pSB229	*GAPDH*	YHE2	17.2	mg/ml	128
	Full-length	-	pAAH5N	*ADH1*	AH22	19	pmol/mg	146
		f	pNW255, pNW249	*GAPDH*	YHE2			**321**
1A2	Full-length	2 ×	pHE36	*GAPDH*	YHE2	7	mg/ml	127, 128
	Full-length	-	PDP34	*GAPDH*	YHE2	191	pmol/mg	194
	Full-length	2 ×	pCS316	*GAPDH*	YHE2	8.4	mg/ml	128
	Full-length	i	pYe DP	*GAL10-CYC1*		N/A	-	132
1B1	N-term. Δ 2-4 aa	-	PYES2	*GAPDH*	JL20	340	pmol/mg	192
2A6	Full-length	-	pAAH5N	*ADH1*	AH22	59	pmol/mg	146
2B6	Full-length	-	pAAH5N	*ADH1*	AH22	95	pmol/mg	146
2C8	Full-length	-	pAAH5N	*ADH1*	AH22	283	pmol/mg	146
2C9	Full-length	-	pAA7	*GAL7*	MT8-1	200	pmol/mg	322
	Full-length	-	pAAH5	*ADH1*	D12	60	pmol/mg	144, 145
	Full-length	-	pAAH5N	*ADH1*	AH22	250	pmol/mg	146
2C18	Full-length	-	pAAH5	*ADH1*	334	60	pmol/mg	323
	Full-length	-	pAAH5N	*ADH1*	AH22	101	pmol/mg	146
2C19	Full-length	-	pAAH5	*ADH1*	334	17	pmol/mg	323
2D6	Full-length	-	pMA91	*PGK*	AH22	50	pmol/mg	152
	Full-length	-	pMA91	*PGK*	AH22	98	pmol/mg	155
	N-term. Δ 1-22 aa	-	pYe DP	*GAL10-CYC1*	2805	160	pmol/mg	324
	N-term. Δ 0.1kb	-	pAAH5	*ADH1*	AH22	32	pmol/mg	146
2E1	Full-length	-	pAAH5N	*ADH1*	AH22	99	pmol/mg	146
3A4	Full-length	-	pAAH5	*ADH1*	AH22	7-11	nmol/L	201
	Full-length	-	pAAH5	*ADH1*	AH22	156	pmol/mg	146
	Full-length	f	pAAH5N	*ADH1*	AH22	52	pmol/mg	129
	Full-length	2 ×	pAAH5N	*ADH1*	AH22	131	pmol/mg	129
	Full-length	f, 2×[3]	pAAH5N	*ADH1*	AH22	85	pmol/mg	129
	Full-length	-	pAAH5	*ADH1*	D12, AH22	N/A	-	325
	Full-length	-	pYeDP1/8	*GAL10-CYC1*	L-41-3	90	pmol/mg	326
	Full-length	i	pAAH5	*GAL10-CYC1*	W303.1B	2-3	nmol/L	327
	Full-length	i	pAAH5	*GAL10-CYC1*	W303.1B	85	nmol/L	132
			pYeDP		AH22, D12	3	nmol/L	**301**
4F2	Full-length	-	pGYR1	GAP	AH22	10.6	nmol/mg	156
4F12	Full-length	i	pYeDP60	*GAL10-CYC1*	W (R)	N/A	-	328
	Full-length	-	pGYR1	GAP	AH22	1.9	nmol/mg	156

[1] Expression of P450s without OR (-) or with OR in certain constructs. f = fusion; 2 × = dual expression system. The P450 and OR cDNAs in a single plasmid are controlled by two independent promoters., i = yeast genomic integration of various expression cassettes habouring the human OR cDNA or human b5 cDNA (132)

[2] pmol/mg = expression level of P450 content per mg of microsomal protein; nmol/L = expression level of nmol P450 enzyme per liter of culture, mg/ml = protein content of microsomes.

[3] Dual expression system of (P450 3A4 cDNA and yeast OR cDNA fused) and human b5 cDNA in one plasmid controlled by two independent promoters.

Stellenbosch University http://scholar.sun.ac.za

17	N-term. Δ 1-18 aa, C-term. 5 × histidine tag	-	pPRL2	*N/A*	YPH259	N/A	-	329
		i	PcDE2, pYeSF	*PGK*	W303B	70	pmol/mg	56
19	Full-length	-	pYeDP1/8	*GAL10-CYC1*	W303B	35	pmol/mg	330
21			pYE8	*GAPDH, ADH*	20B-12	0.8	nmol/mg	331
Bovine								
2C8			pAAH5	*Q*	334	250	pmol/mg	323
17	Full-length	-	pAAH5		SEY6210	2.5	nmol/L	88
	Full-length	f	pAAH5	*ADH1*	AH22	N/A	-	123
	Varous lengths	f	pAAH5	*ADH1*	AH22	N/A	-	130
	-	-	YEp51	*GAL10*	GRF18	37	mg/L	202
21	Full-length	f	pAAh5	*ADH1*	AH22	N/A	-	126
Pig								
4A25, 4A26	Full-length	i	pYeDP60	*GAL10-CYC1*	W (R)	N/A	-	332
Rat								
1A1	Full-length	-	pAAH5	*ADH1*		1.4	%[4]	**101**
	Full-length	f	pAAH5	*ADH1*	AH22	N/A	-	124
	Full-length	f	pAFCR!	*ADH1*	AH22	110 10	pmol/mg pmol/mg◆	122
	Full-length	f	pAMR2	*ADH1*	AH22	270 240	pmol/mg pmol/mg◆	125
1A2	Full-length	-	pAM82	*PHO5*	AH22	1.0	%[5]	147, 148
2A1			pAAH5					333
2B1			pmA56	*ADCI*		0.25 0.15	% nmol/mg	**N**
2B2			pAM82	*PHO5*		30	pmol/mg	**000 1**
2C11	Full-length	-	pYcDE2	*ADH1*	TD1	100	pmol/mg	334
4A1			pAAH5	*ADH1*				201
7B								
27	Replacement of target signal with bovine P450 17	2 ×[6]	pRXMS25	*PGK*	AH22	150	pmol/mg	153, 154
Mouse								
1A1	Full-length	-	pAAH5	*ADH1*	50.L4	N/A	-	335
	Full-length	-	pYe DP1,10	*GAL10-CYC1, PGK*	W303.1B	0.6	%[4]	191
			pYe DP1,10		W303.1B			120
	-	i[7]	pRS404	*ADH1*	YPH500	38.8	pmol/mg	136
Rabbit								
1A1, 1A2	Full-length	-	pYe DP1,10	*GAL10-CYC1, PGK*	W303.1B	0.26 0.7	%[4]	336, 337
2B4			YEp51	*GAL10*	334			338
2B5			YEp51	*GAL10*	334			338
2C14			pAAH5	*ADH1*		1.4	%	**T,u**
P450 (ω-1) P450(16α)	-	-	pAAH5	*ADH1*	AH22	N/A	-	339
P450 pAHF3			pAAH5					**W**
2E1	Full-length	-	pAAH5, YEp13	*CUP1*	50.L4	30-60	pmol/mg	149
2E2	Full-length	-	YEp13	*CUP1*	50.L4	70	pmol/mg	150

---

[4] Expression level given in % (w/w) of the protein relative to the total proteins in yeast microsomal fractions.

◆ the content of NADPH-P450 reductase in microsomes

[5] estimated percentage of expressed P450 of the total yeast protein

[6] Co-expression of P450c27 with bovine adrenodoxin and NADPH-adrenodoxin reductase

[7] i = yeast genomic *Ty* transposon integration of multi-integration (six copies) expression cassettes with the hamster OR cDNA

Stellenbosch University http://scholar.sun.ac.za

P450 ka$_1$, P450 ka$_2$	Full-length	-	pAAH5	*ADH1*	AH22	60-180	pmol/mg	340
LM4, LM6			pYe DP1,10	*GAL10-CYC1*	W303.1B			337
Canine								
2B11			YEp51	*GAL10*	334			338
Yeast								
52A3	N-term. truncated 47 or 97aa		YEp51	*GAL10*	GRF18	N/A	-	341
	Full-length, insert Xa-factor recognition site at 62-66 aa	-	YEp51	*GAL10*	GRF18	620	pmol/mg	221
	N-term. truncated 66 aa	-	YEp51	*GAL10*	GRF18	340 - 750	pmol/mg	342
Plant								
71A10	Full-length	i	PYeDP60	*GAL10-CYC1*	W (R)	N/A	-	343
71B7			pGY2	*PGK*	ATAH5	N/A	-	121
71D10	Full-length		PYeDP60	*GAL10-CYC1*	W (R)	N/A	-	343
71D13	Full-length	i	PYeDP60	*GAL10-CYC1*	WAT11U, WAT21U	80	pmol/mg	298
71D15	Full-length	i	PYeDP60	*GAL10-CYC1*	WAT11U, WAT21U	50	pmol/mg	298
71D18	Full-length	i	PYeDP60	*GAL10-CYC1*	WAT11U, WAT21U	527	pmol/mg	298
73A	Full-length	i	PYeDP60	*GAL10-CYC1*	W303.1B	200	pmol/mg	344 345
73A17	N-term. modification 36 aa	i	PYeDP60	*GAL10-CYC1*	W303.1B	300	pmol/mg	189
77A3	Full-length	i	PYeDP60	*GAL10-CYC1*	W (R)	252	pmol/mg	343
78A3	Full-length	i	PYeDP60	*GAL10-CYC1*	W (R)	N/A	-	343
82C1	Full-length	i	PYeDP60	*GAL10-CYC1*	W (R)	N/A	-	343
83A1, 83B1								
83D1	Full-length	i	PYeDP60	*GAL10-CYC1*	W (R)	11	pmol/mg	343
93C1	Full-length	i	PYeDP60	*GAL10-CYC1*	W (R)	N/A	-	343
97B2	Full-length	i	PYeDP60	*GAL10-CYC1*	W (R)	N/A	-	343
98A2	Full-length	i	PYeDP60	*GAL10-CYC1*	W (R)	N/A	-	343
P450 taxane			PYeD60	*GAL10-CYC1*	WAT11, WAT 21			346
Insect								
6D1			pYES2	*GAL10-CYC1*		0.18	nmol/mg	347

Stellenbosch University http://scholar.sun.ac.za

Table 2.2 D

Summary of heterologous expression of cytochromes P450 in *P. pastoris*.

CYP	Modification	Vector	Promoter	Strain	Expression level[1]		Ref.
Human:							
17	Full-length C-term 6×histidine tag	pPIC3.5K	*AOX1*	GS115	300	pmol/mg	165
Crustacean							
2L1	Full-length	pPICZa	*AOX1*	GS115	20	pmol/mg	348
Shark							
17	Full-length	pPHIL	*AOX1*	GS115	N/a	-	349
Plant							
79D1	Full-length Two silent mutaions at codon 2 and 6	pPICZc	*AOX1*	X-33	128	pmol/mg	164

[1] pmol/mg = Expression level of P450 content per mg of microsomal protein.

Stellenbosch University http://scholar.sun.ac.za

SECTION 2.3: Heterologous protein secretion by yeasts' secretory pathway

2.3.1 Introduction

The eukaryotic secretory pathway plays an essential role in maintaining the cellular requirements for biochemical compartmentation and it is mediated by biochemically and morphologically distinct membrane enclosed organelles (350), shown in Figure 2.13. In these organelles activities such as protein transport, sorting of proteins from the cytoplasm to various intracellular organelles and cell surface transport takes place. Newly expressed proteins have to pass through this multi-component secretory apparatus, within which disulphide bond formation, proteolytic maturation of prohormones, N- and O-linked glycosylation and other post-translational modifications occur (351). The synthesised proteins in the cytoplasmic ribosomes are inserted into the lumen of the ER and are subsequently transported to the Golgi complex. From there, the protein is delivered to the cell surface via the Golgi complex-derived secretory vesicle and the plasma membrane.

In the field of biotechnology, the potential of exploiting this secretory pathway to produce recombinant proteins was soon realised and many proteins were successfully secreted (352). The secretion of heterologous proteins from yeast has many advantages over intracellular expression. Intracellular expressed proteins are mostly insoluble and may not adopt their correct conformation (e.g. disulfide bonding) in the cytosol. This can be prevented with extracellular expression, since the secretion of foreign proteins undergo proper posttranslational processes within the secretory compartments ensuring correctly folded proteins (353). Secretion is highly advantageous because of the initial purity of the product in the substantially protein-free culture medium leading to a minimal purification from the culture medium. The choice between intracellular or extracellular expression also lies in the type of protein that needs to be expressed. Some protein products are toxic or unstable when produced cytoplasmically, which can be avoided when they are secreted into the medium (354).

Stellenbosch University http://scholar.sun.ac.za

In this Chapter we look at each of the individual secretory compartments of *S. cerevisiae* and describe their role and function. Secretion of proteins in yeasts is a very complex process and the exact detail of this is beyond the scope of this thesis. Therefore, a basic outline of the secretory pathway is described in this chapter and excellent comprehensive reviews are to be found in (350, 355, 356, 357, 358, 359). This section discusses some of the factors involved in secretion of heterologous proteins from *S. cerevisiae*. Additionally, we outline the potential limitations of secreting expressed cytochrome P450 proteins with regards to hydrophobicity, glycosylation and size.

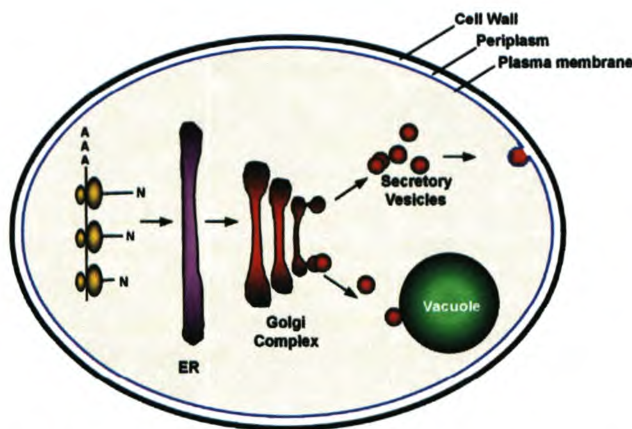

Figure 2.13 An illustration of the yeast secretory pathway. Yeast secretory glycoproteins synthesised in the cytoplasm are transported over the ER membrane to the ER lumen where they undergo N-linked core glycosylation and correct folding. After passage through the Golgi complex, where the outer chain glycosylation modifies the secretory protein, the glycosylated proteins are delivered either to the vacuole or the cell surface via secretory vesicles.

2.3.1.1 Passage of protein from the cytoplasm to ER and processing in the ER

There are four distinct events that occur during the protein passage from the cytoplasm to the ER and in the lumen of the ER — (a) the initial sorting event, where proteins destined for export are recognised by their amino-terminal signal peptides, which consists of a precursor form, (b) the delivery of the protein with its secretory precursor polypeptide to the target membrane of the ER, where translocation of the protein through the proteinaceous pores of the membrane occurs, (c) the proteolytic processing of the signal peptide and (d) the correct folding and N-glycosylation of the secreted protein.

Stellenbosch University http://scholar.sun.ac.za

2.3.1.1.1 Leader sequences for initiation of secretion

For a protein to enter the *S. cerevisiae* secretory pathway, it requires an amino-terminal hydrophobic signal peptide. There are two classes of signal sequences — one uses a heterologous signal peptide, often derived from the protein being secreted, and the other uses a homologous yeast signal peptide. It was found that the majority of heterologous signal peptides will not direct efficient secretion of a protein from *S. cerevisiae* (360), and subsequent research endeavours concentrated on the more reliable homologous *S. cerevisiae* signal sequences. The most frequently used homologous signal sequences are the yeast invertase (SUC2) (361, 362, 353), acid phosphatase (PHO2) and prepro-α-factor (MFα1) (363, 364). Only the latter will further be discussed, as it is the most successfully used signal sequence and it was used in our expression experiments.

The yeast pheromone α-factor is a tridecapeptide mating factor, 12-13 L-amino acid residues in length, produced and secreted by haploid cells of the α-mating type and is required for efficient mating of haploid α and a cells to form diploids in the G1 phase of the cell division cycle. The α-factors are derived by proteolytic processing from a larger precursor, the prepro-α-factor, which is encoded by the MFα1 gene. The prepro-α-factor protein consists of the pre region, a hydrophobic signal sequence of 22 amino acids, a pro region of approximately 61 amino acids and the mature α-factor (mating factor and pheromone) sequence of four tandem repeats of 13 amino acids seen in Figure 2.14. The repeats are flanked by short peptides consisting of Lys-Arg-(Glu/Asp-Ala)$_{2-3}$ (365). The pre region of the prepro-α-factor is the actual signal peptide that acts as the specific target signal and has three distinct domains which is common in most signal peptides: (i) a positively charged amino-terminus (n-region, 1-5 residues), (ii) a central, hydrophobic part (h-region, 7-15 residues) and (iii) a more polar carboxyl-terminus (c-region, 3-7 residues) (366, 367).

Signal sequences can direct proteins to a membrane through different targeting pathways and select different translocation systems such as the signal recognition

Stellenbosch University http://scholar.sun.ac.za

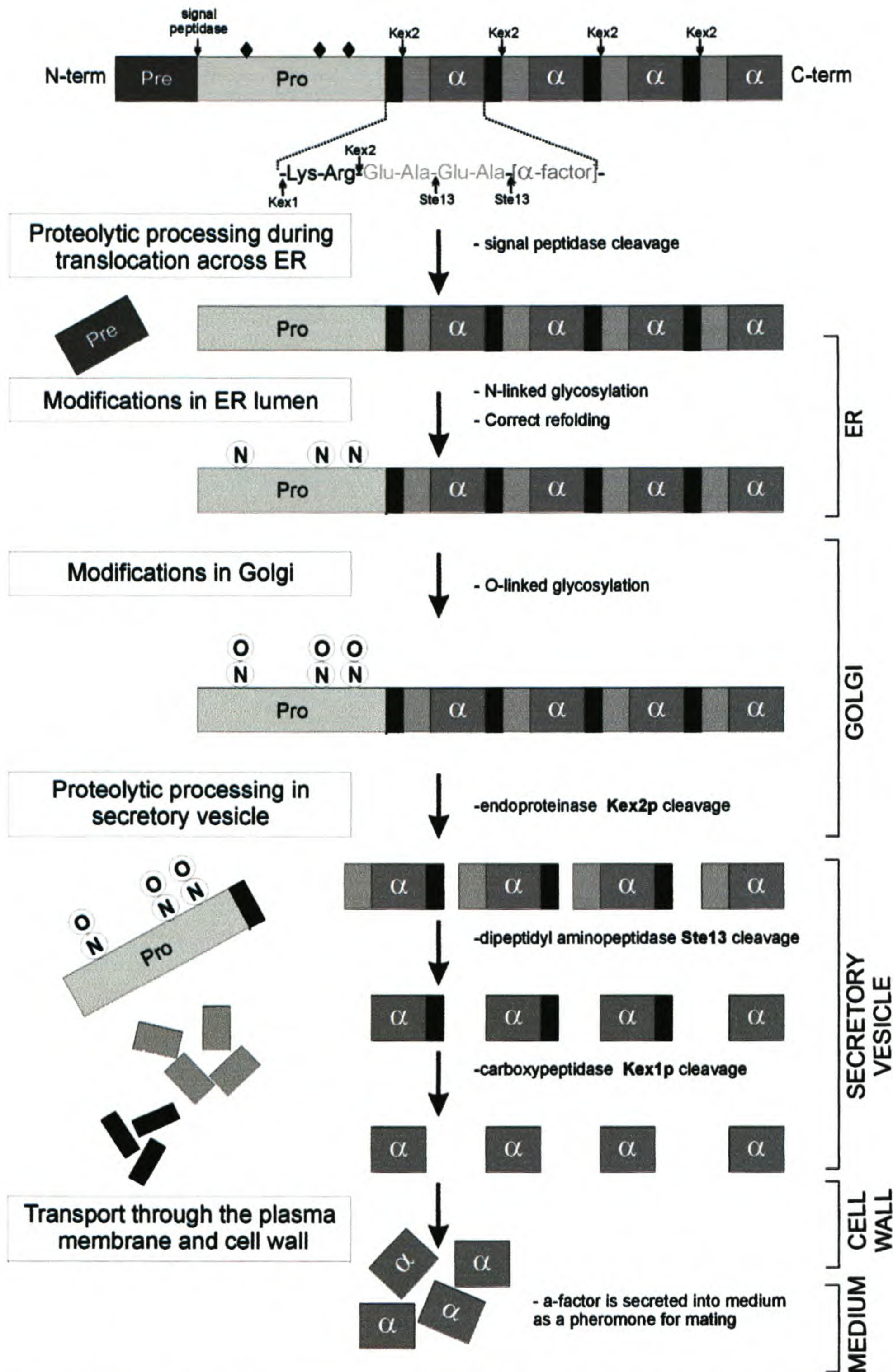

Figure 2.14 The processing of the prepro-α-factor in the secretory pathway. Signal peptidase in the ER cleaves the signal peptide, Kex2p endoprotease in the Golgi cleaves at carboxyl terminal of Lys-Arg, Ste13 diaminopeptidase removes the Glu-Ala pairs, Kex1p protease removes Lys and Arg at carboxyl terminus. N-and O-linked glycosylation of the pro-region at three sites is shown as **N** and **O**, respectively.

partical (SRP)-dependent or SRP-independent pathways (368). Studies in yeast have revealed that the hydrophobic region (h-region) of a signal sequence is the parameter

Stellenbosch University http://scholar.sun.ac.za

discriminating between SRP-dependent and SRP-independent pathways (369). The prepro-α-factor precursor uses the SRP-dependent pathway.

2.3.1.1.2 The translocation of the protein from the cytoplasm to the ER

The signal peptide, usually located at the amino-terminus of nascent polypeptides, initiates the cotranslational export of the secretory protein from the cytoplasm into the ER by binding to the SRP. The SRP is a ribonucleoprotein particle consisting of a 300-nucleotide RNA component (7SL RNA) which is complexed with six unique polypeptides of 9, 14, 19, 54, 68 and 72 kDa (370). It serves as a cytoplasmic adapter between nascent secretory polypeptide chains and the protein translocation apparatus of the ER membrane, shown in Figure 2.15. After the SRP binds to the ribosome carrying the nascent polypeptide chain with the signal sequence, the complex is targeted to the ER membrane. The targeting process is regulated by a nascent-polypeptide-associated complex (NAC) (371), which has been implicated in modulating the interaction between the ribosome and SRP (372). At the ER membrane the complex binds to the SRP receptor, a heterodimeric integral membrane protein, located in the ER membrane.

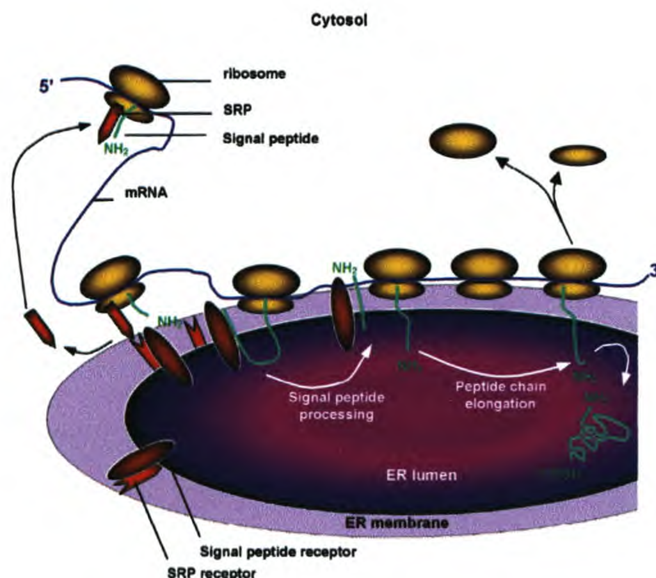

Figure 2.15 A schematic illustration of the export of a secretory protein from the cytoplasm into the ER (signal hypothesis). Nascent polypeptide chains are bound by a signal-recognition particle (SRP) to form a soluble SRP-nascent chain-ribosome complex. This is directed to the ER membrane through the interaction of the SRP with a heterodimeric membrane-bound SRP receptor (docking protein). The signal sequence receptor is engaged by the signal peptide resulting in a release of the SRP into the cytosol ready for another round of ER targeting. The mechanism of translocation of the nascent polypeptide is unknown, the signal-peptide cleavage occurs either during or immediately after the translocation step.

Stellenbosch University http://scholar.sun.ac.za

The nascent chain-ribosome complex associates with the protein translocation channel after the targeting step. This channel provides an aqueous environment in which the hydrophilic nascent secretory peptide chain can be translocated (373).

2.3.1.1.3 The processing of the prepro-α-factor during translocation

The translocation of the secretory protein into the ER is initiated by the hydrophobicity of the pre region, directed by the signal sequence. The pre region is cotranslationally cleaved by a highly specific signal peptidase (Figure 2.14) resulting in a pro-α-factor.

2.3.1.1.4 Correct folding of heterologous proteins and N- glycosylation

After the heterologous protein with its modified α-factor secretion signal peptide has been translocated into the lumen of the ER, it needs to undergo correct folding and N-glycosylation prior to its continuation through the secretory pathway.

Correct folding of the secretory protein is facilitated by a number of lumenal proteins. They include the Kar2p/BiP (a chaperone) and the protein disulfide isomerase (PDI), which mediates the formation of disulfide bonds (374). The efficiency and accuracy of the yeast's folding mechanism, which uses similar components and machinery as the mammalian ER, produces authentic refolded mammalian heterologous proteins (375). Efficient secretion of the heterologous protein will only be achieved once the protein is correctly folded.

Nascent polypeptides, translocating across the ER membrane, are cotranslational (i.e. membrane translocation is obligatorily coupled with protein synthesis) or posttranslationally (i.e. membrane translocation follows after protein synthesis) N-glycosylated by an oligosaccharyltransferase complex (OST) at the potential glycosylation motif (Asn-X-Thr/Ser) (376), as seen in Figure 2.16. Residue X can be any amino acid except proline. The inner core of the glycosylated chain consists of three glucose (Glc), nine mannose (M) and two N-acetylglucosamine (GlcNAc) residues, which are linked to the N-amide group of Asn, hence N-linked glycosylation. The pro region of the MFα1 factor is also glysosylated at three potential glycosylation sites

Stellenbosch University http://scholar.sun.ac.za

(377). Investigations into the glycosylation of the precursor structure showed that N glycosylation of pro-α-factor is important, but not essential for the transport of the ER to the Golgi. Mutations of any of the three glycosylation sites in the pro region resulted in a slight reduction in α-factor secretion and intracellular accumulation of the pro-α-factor (378). Therefore it was concluded N-linked oligosaccharides increase the solubility and stability of many proteins, ensuring correct folding of secretory proteins.

Figure 2.16 Typical structure of yeast N-and O-linked oligosaccharides. Asparagine-linked core oligosaccharides **(A)** are assembled and added in the ER. N-linked core oligosaccharides are usually modified during transit through the Golgi complex by the addition of an outer chain carbohydrate. **(B)** M: mannose, P: phosphate; GlcNAc: *N*-acetylglucosamine; Glc: glucose. α-1,6; α-1,3; α-1,2 and β-1,4 linkages are designated.

2.3.1.2 Transport from the ER to the Golgi complex

There are three steps involved in the transport from the ER to the Golgi complex – a) budding of transport vesicles, COPI-, COPII- and clathrin-coated vesicles, from the cell surface; b) targeting of the ER-derived vesicle to the Golgi complex and c) fusion of the transport vesicle to the acceptor membrane. A detailed biochemical dissection of these processes has been well documented and models have been postulated (357, 379, 380). Three GTP-binding proteins Arr1p, Arf1p and Ypt1p have been implicated in transport from the ER to the Golgi complex in *S. cerevisiae* (356, 381,).

Stellenbosch University http://scholar.sun.ac.za

2.3.1.3 Golgi processing and passage from Golgi to secretory vesicles

The Golgi complex plays an important role in regulating several aspects of protein and membrane movement throughout the cell. This organelle is the compartment for sorting proteins from the constitutive secretory pathway to the lysosomal and regulated secretory pathways. Additional functions are the regulation of membrane recycling from the cell surface and involvement with the endocytic pathway (382). Although the complex functions of this organelle remain unclear, two functions executed by the Golgi apparatus, namely outer chain glycosylation and proteolytic processing of the α-factor, are of interest for our investigations into the extracellular heterologous protein expression and will therefore be addressed.

2.3.1.3.1 Outer chain glycosylation

In the Cis-Golgi phase, the ER-secreted protein or pro-α-factor with its inner core is elongated by the addition of an outer chain containing 50 or more mannose units (Figure 2.16). The outer-chain carbohydrates consists of α-D-mannose residues in α-(1→6), α-(1→2) and α-(1→3) linkages together with a limited number of mannobiosylphosphate groups in diester linkage. The order of glycosylation from the core oligosaccharide to the matured N-linked glycans and the compartmentalisation of mannosyltransferase activities are well documented (383, 384).

This elongation by the addition of mannose residues of the pro-α-factor outer chain in *S. cerevisiae* (hyperglycosylation) does not occur in higher eukaryotes. In mammalian Golgi cells, mannosidases trim back core oligosaccharides to three mannose residues and substitute the residues with sialic acid, galactose and N-acetylglucosamine (385). These significant differences in the mechanism of glycosylation between the yeast-derived and native glycoproteins consequently limit the use of *S. cerevisiae* as a secretory expression host. The addition of large numbers of non-native sugar residues to heterologous proteins, can also act as immunogenic determinants (386).

Stellenbosch University http://scholar.sun.ac.za

2.3.1.3.2 Proteolytic processing step

In the trans-Golgi phase an endoprotease, Kex2p, cleaves the dibasic residues, Lys-Arg, at the carboxyl-terminus, separating the pro-region from the α-factor. During the formation of the secretory vesicle, destined for the plasma membrane, dipeptidyl aminopeptidase A, Ste13p, remove the dipeptides Glu-Ala, Asp-Ala or Val-Ala from the amino-terminus of the α-factor. Carboxypeptidase B (Kex1p) removes Lys-Arg from the carboxyl-termini of the first three excised units to generate mature α-factor peptides (Figure 2.14).

2.3.1.4 Passage of the Golgi complex-derived vesicles and fusion to the plasma membrane

At least two different classes of secretory vesicles, which bud from the trans-Golgi and fuse to the plasma membrane, have been identified (387). Fusion to the plasma membrane requires vesicle- and target membrane specific factors, such as the SNARE (soluble N-ethylmaleinide sensitive factor attachment protein receptor) complex (388) and the GTPase protein complex, called Sec4p (389).

The accepted model for the function and interaction of the Sec4p protein in the fusion of secretory vesicles is illustrated in Figure 2.17. The soluble Sec4p exchanges GDP for GTP. The Sec4p-GTP complex attaches to a newly formed secretory vesicle and moves towards the plasma membrane. The docking and fusion onto the plasma membrane is facilitated by a membrane bound effector protein which recognise the Sec4p protein. After the release of the proteins from the vesicle into the periplasmic space, the Sec4p-GTP vesicle complex dissociates from the membrane due to a conformational chance induced by GTP hydrolysis. The Sec4p-GDP complex is then ready for another cycle of secretory vesicle docking and fusion.

Stellenbosch University http://scholar.sun.ac.za

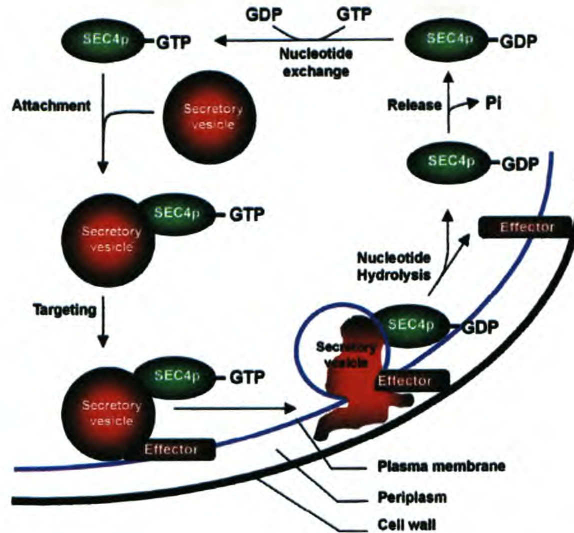

Figure 2.17 Schematic illustration of the passage and fusion of the secretory vesicle to the plasma membrane. Soluble Sec4p exchanges GDP for GTP and binds to a newly formed secretory vesicle, resulting in the interaction with a plasma membrane effector protein. Engagement of the vesicle with the effector protein induces the fusion event and the recombinant proteins inside the secretory vesicle moves into the periplasm. Sec4p-GDP disengages from the effector protein by guanine nucleotide hydrolysis so that both proteins can then be reutilised for another round of secretory targeting and fusion. Adapted from (390).

2.3.1.5 Passage of proteins through cell wall

Whether a secreted protein is retained in the periplasmic space or cell wall, or diffuses into the culture medium depends on the characteristics of the secreted protein (e.g. molecular size, shape, electric charge and glycosylation (391, 392) and is also highly dependable on the cell wall permeability which is influenced by the following factors. Firstly, the permeability varies depending on the physiological condition the yeast cells were subjected to. Fast growing cells have higher cell wall porosity than slow-growing cells due to stretching of the cell wall (393), thereby increasing permeability. Secondly, the permeability of the cell wall is also dependent on the constituency of the mannoproteins in the cell wall which limit the porosity due to disulphide bonds, ionic interactions and large mannan side-chains (394, 395). The mannoproteins in the cell wall differ from strain to strain (396). Finally, the composition of the medium also affects the cell wall porosity (394).

Stellenbosch University http://scholar.sun.ac.za

2.3.2 Discussion

The secretory pathway in eukaryotes is a highly complex system that can be summarised as follows — (i) proteins which enter the secretory pathway are specified by short peptide sequences within the protein which determine specific interactions with particular components of the secretory machinery; (ii) proteins are transported between organelles in vesicles, which bud from a donor organelle upon assembly of cytosolic protein complexes and which subsequently fuse with an acceptor organelle upon assembly of complexes composed of proteins from both vesicle and acceptor membranes and (iii) the macromolecular assembly along the secretory pathway is coupled to nucleotide hydrolysis by the action of either ATPases or GTPases. New genes that encode specific proteins involved in the secretory pathway are discovered daily and new insight will lead to improved systems for the heterologous expression and production of extracellular proteins.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 3: GOALS OF THE STUDY

- To update and compare heterologous expression systems used for the expression of P450 enzymes.
- To find suitable expression systems that produce full-length steroidogenic P450 enzymes at large-scale.
- To determine and suggest cloning strategies that increase expression levels.
- To overexpress unmodified human P45017α, b_5 and aromatase in a heterologous expression system that will secrete the protein into the medium.
- To overexpress unmodified human P45017α intracellularly.
- To develop methods for preparation of yeast microsomes containing expressed P450 enzymes.
- To develop a purification protocol that will be rapid and cost efficient.
- To develop purification steps that will yield ultrapure recombinant P450 enzymes for future antibody production and crystallisation.

REFERENCES

1 Beresford, A. P., Taylor, R. J., Ashcroft, J.-A., Ayrton, J., Tucker, G. T. and Ellis, S. W. (1996) *Xenobiotica* **26**, 1013-1023.

2 Lu, A. Y. H. (1998) *Drug Metab. Dispos.* **26**, 1168-1173.

3 Hall, P. F. (1985) *Vitam. Horm.* **42**, 315-368.

4 Omura, T. and Sato, R. (1962) *J. Biol. Chem.* **237**, 1375-1376.

5 Omura, T. and Sato, R. (1964) *J. Biol. Chem.* **239**, 2370-2385.

6 Hall, P. F. (1985) *Ann. NY. Acad. Sci.* **458**, 203-215.

7 Estabrook, R. W., Shet, M. S., Fisher, C. W., Jenkins, C. M. and Waterman, M. R. (1996) *Arch. Biochem. Biophys.* **333**, 308-315.

8 Lewis, D. F. V. and Pratt, J. M. (1998) *Drug Metabol. Rev.* **30**, 739-786.

9 Hlavica, P. and Lewis D. F. V. (2001) *Eur. J. Biochem.* **268**, 4817-4832.

10 Coon, M. J., Vaz, A. D. N., McGinnity, D. F. and Peng, H.-M. (1998) *Drug Metab. Dispos.* **26**, 1190-1193.

11 Newcomb, M., Shen, R. and Choi, S.-Y. (2000) *Am. Chem. Soc.* **122**, 2677-2686.

12 Guengerich, F. P. (2001) *Chem. Res. Toxicol.* **14**, 611-650.

13 Lewis, D. F. V. (2003) *Pharmacogenomics* **4**, 387-395.

14 Fulco, A.J. (1991) *Annu. Rev. Pharmacol.* **31**, 177-203.

15 Lambeth, J. D., Seybert, D. W., Lancaster, J. R., Salerno, J. C. and Kamin, H. (1982) *Mol. Cell Biochem.* **45**, 13-31.

16 Ortiz de Montellano, P. R. (1986) 556 pp. In: *Cytochrome P450: Structure, Mechanism and Biochemistry*. Paul R. Ortiz de Montellano (Ed), Plenum Press, New York.

17 Miwa, G. T., West, S. B., Huang, M. T. and Lu, A. Y. (1979) *J. Biol. Chem.* **254**, 5695-5700.

18 Shen, A. L. and Kasper, C. B. (1995) *J. Biol. Chem.* **270**, 27475-27480.

19 Sakaguchi, M., Mihara, K and Sato, R. (1987) *EMBO J.* **6**, 2425-2431.

20 Szczesna-Skorupa, E. and Kemper, B. (1989) *J. Cell. Biol.* **108**, 1237-1243.

21 Roux, S. P. (1995) M. Sc. thesis, University of Stellenboch.

22 Black, S.D. (1992) *FASEB J.* **6**, 680-685.

23 Sato, T., Sakaguchi, M., Mihara, K. and Omura, T. (1990) *EMBO J.* **9**, 2391-2397.

24 Nelson, D. R. and Strobel, H. W. (1988) *J. Biol. Chem.* **263**, 6038-6050.

25 Kuroda, R., Kinoshita, J., Honosho, M., Mitoma, J. and Ito, A. (1996) *J. Biochem.* **120**, 828-833.

Stellenbosch University http://scholar.sun.ac.za

26 Tanaka, S., Kinoshita J.-y., Kuroda, R. and Ito, A. (2003) *J. Biochem.(Tokyo)* **133**, 247-251.

27 Miller, W. L. (1988) *Endocr. Rev.* **9**, 295-318.

28 Stocco, D. M. and Clark, B. J. (1996) *Endocr. Rev.* **17**, 221-244.

29 Bose, H. S., Lingappa, V. R. and Miller, W. L. (2002) *Nature* **417**, 87-91.

30 Waterman, M. R. and Simpson, E. R. (1985) *Mol. Cell. Endocrinol.* **39**, 81-89.

31 Conley, A. J. and Bird, I. M. (1997) *Biol. Reprod.* **56**, 789-799.

32 Katagiri, M., Suhara, K., Shiroo, M. and Fujimura, Y. (1982) *Biochem. Biophys. Res. Comm.* **108**, 379-384.

33 Katagiri, M., Kagawa, N. and Waterman, M. R. (1995) *Arch. Biochem. Biophys.* **317**, 343-347.

34 Fevold, H. R., Lorence, M. C., McCarthy, J. L., Trant, J. M., Kagimoto, M., Waterman, M. R. and Mason, J. I. (1989) *Mol. Endocrinol.* **3**, 968-975.

35 Cloutier, M., Fleury, A., Courtemanche, J., Ducharme, L., Mason, J. I. and LeHoux, J.-G. (1995) *Ann. N. Y. Acad. Sci.* **774**, 294-296.

36 Sakai, Y., Tanaka, M., Avdachi, S., Miller, W. L. and Nagahama, Y. (1992) *FEBS Lett.* **301**, 60-64.

37 Trant. J. M. (1995) *Gen. Comp. Endocr.* **102**, 173-182.

38 Tremblay, Y., Fleury, A., Beaudoin, C., Vallee, M. and Belanger, A. (1994) *DNA Cell Biol.* **13**, 1199-1212.

39 Fevold, J. D., Ivanovitch, J. D., Zanger, U. M. and Waterman, M. R. (1993) *Mol. Cell. Endocrinol.* **95**, 95-100.

40 Lin, D., Black, S. M., Nagahama, Y. and Miller, W. L. (1993) *Endocrinology* **132**, 2498-2506.

41 Swart, A. C., Kolar, N. W., Lombard, N., Mason, J. I. and Swart, P. (2002) *Eur. J. Biochem.* **269**, 5608-5616.

42 LeHoux, J.-G., Mason, J. I. and Ducharme, L. (1992) *Endocrinology* **131**, 1874-1882.

43 Brière, N., Martel, D., Cloutier, M. and LeHoux, J.-G. (1997) *J. Histochem. Cytochem.* **45**, 1409-1416.

44 Mathieu, A. P., Auchus, R. J. and LeHoux, J.-G. (2002) *J. Steroid Biochem.* **80**, 99-107.

45 Van Weerden, W. M., Bierings, H. G., van Steenbrugge, G. J., de Jong, F. H. and Schroder, F. H. (1992) *Life Sci.* **50**, 857-861.

46 Swart, P., Swart, A. C., Waterman, M. R., Estabrook, R. W. and Mason, J. I. (1993) *J. Clin. Endocr. Metab.* **77**, 98-102.

47 Swinney, D. C. and Mak, A. Y. (1994) *J. Am. Chem. Soc.* **114**, 8309-8310.

48 Lee-Robichaud, P., Wright, J. N., Akhtar, M. and Akhtar, M. (1995) *Biochem J.* **308**, 901-908.

49 Miller, W.L., Auchus, J. and Geller, D.H. (1997) *Steroids* **62**, 133-142.

50 Geller, D. H., Auchus, R. J., Mendonça, B. B. and Miller, W. L. (1997) *Nature Genet.* **17**, 201-205.

51 Geller, D. H., Auchus, R. J. and Miller, W. L. (1999) *Mol. Endocrinol.* **13**, 167-175.

52 Zhang, L., Rodriguez, H., Ohno, S. and Miller, W. L. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 10619-10623.

53 Miller, W. L and Auchus, R. J. (2000) *J. Clin. Endocr. Metab.* **85**, 1346.

54 Kominami, S., Ogawa, N., Morimune, R., Huang, D. Y. and Takemori, S. (1992) *J. Steroid Biochem. Mol. Biol.* **42**, 57-64.

55 Swart, P., Engelbrecht, Y., Bellstedt, D. U., de Villiers, C. A. and Dreesbeimdieke, C. (1995) *Endocr. Res.* **21**, 297-306.

56 Auchus, R. J., Lee, T. C. and Miller, W. L. (1998) *J. Biol. Chem.* **273**, 3158-3165.

57 Yamazaki, H., Johnson, W. W., Ueng, Y. F., Shimada, T and Guengerich, F. P. (1996) *J. Biol. Chem.* **271**, 27438-27444.

58 Lee-Robichaud, P., Kaderbhai, M. A., Kaderbhai, N., Wright, J. N. and Akhtar, M. (1997) *Biochem. J.* **321**, 857-863.

59 Hasemann, C. A., Kurumbail, R. G., Boddupalli, S. S., Peterson, J. A. and Deisenhofer, J. (1995) *Structure* **3**, 41-62.

60 Shozu, M. and Simpson, E. (1998) *Mol. Cell. Endocrinol.* **139**, 117-129.

61 Bulun S. E., Takayama, K., Suzuki, T., Sasano, H., Yilmaz, B. and Sebastian, S. (2004) *Semin. Reprod. Med.* **22**, 5-9.

62 Simpson, E., Rubin, G., Clyne, C., Robertson, K., O'Donnell, L., Jones, M. and Davis, S. (2000) *Trends Endocrin. Met.* **11**, 184-188.

63 Toma, Y., Higashiyama, T., Yarborough, C. and Osawa Y (1996) *Endocrinology* **137**, 3791-3796.

64 Graham-Lorence, S., Amarneh, B., White, R. E., Peterson, J. A. and Simpson, E. R. (1995) *Protein Sci.* **4**, 1065-1080.

65 Matteson, K. J., Picado-Leonard, J., Chung, B.-C., Mohandas, T. K. and Miller, W. L. (1986) *J. Clin. Endocr. Metab.* **63**, 789-791.

66 Nebert, D. W., Nelson, D.R., Coon, M. J., Estabrook, R. W., Feyereisen, R., Fujii-Kuriyama, Y., Gonzalez, F. J., Guengerich, F. P., Gunsalus, I. C., Johnson, E. F., Loper, J. C., Sato, R., Waterman, M. R. and Waxman, D. J. (1991) *DNA* **10**, 1-14.

67 Zuber, M. X., John, M. E., Okamura, T., Simpson, E. R. and Waterman, M. R. (1986) *J. Biol. Chem.* **261**, 2475-2482.

Stellenbosch University http://scholar.sun.ac.za

68 Graham-Lorence, S., Khalil, M. W., Lorence, M. C., Mendelson, C. R. and Simpson, E. R. (1991) *J. Biol. Chem.* **266**, 11939-11946.

69 Corbin, C. J., Graham-Lorence, S., McPhaul, M., Mason, J. I., Mendelson, C. R. and Simpson, E. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8948-8952.

70 Amarneh, B., Corbin, C. J., Peterson, J. A., Simpson, E. R. and Graham-Lorence, S. (1993) *Mol. Endocrinol.* **7**, 1617-1624.

71 Sethumadhavan, K., Bellino, F. L. and Thotakura, N. R. (1991) *Mol. Cell. Endocrinol.* **78**, 25-32.

72 Chen, S., Zhou, D., Swiderek, K. M., Kadohama, N., Osawa, Y. and Hall, P. F. (1993) *J. Steroid Biochem. Mol. Biol.* **44**, 347-350.

73 Amarneh, B., Corbin, C. J., Peterson, J. A., Simpson, E. R. and Graham-Lorence, S. (1993) *Mol. Endocrinol.* **7**, 1617-1624.

74 Nestler, J.E. and Jakubowicz, D. J. (1996) *New Eng. J. Med.* **335**, 617-623.

75 Qin, K. and Rosenfield, R. L. (1998) *Mol. Cell. Endocrinol.* **145**, 111-121.

76 Miller, W. L. (1997) *J. Mol. Endocrinol.* **19**, 227-240.

77 Sasano, H. and Harada, N. (1998) *Endocr. Rev.* **19**, 593-607.

78 Zuber, M. X., Simpson, E. R. and Waterman, M. R. (1986) *Science* **234**, 1258-1261.

79 Mathew, P. A., Mason, J. I., Trant, J. M. and Waterman, M. R. (1990) *Mol. Cell. Endocrinol.* **73**, 73-80.

80 Lindberg, L. P. and Negishi, M. (1989) *Nature* **339**, 632-634.

81 Gonzalez, F. J. and Korzekwa, K. R. (1995) *Annu. Rev. Pharmacol. Toxicol.* **35**, 369-390.

82 Gasser, R., Funk, C., Matzinger, P., Klemisch, W. and Viger-Chougnet, A. (1999) *Toxicol. In Vitro* **13**, 625-632.

83 Fujita, K-i. and Kamataki, T. (2002) *Drug Metabol. Pharmacokin.* **17**, 1-22.

84 Their, R., Pemble, S., Kramer, H., Taylor, J. B. and Guengerich, F. P. (1996) *Carcinogenesis* **17**, 163-166.

85 Kushida, H., Fujita, K., Suzuki, A., Yamada, M., Nohmi, T. and Kamataki, T. (2000) *Mutat. Res.* **471**, 135-143.

86 Shimada, T., Oda, Y., Gillam, E. M. J., Guengerich, F. P. and Inoue, K. (2001) *Drug Metab. Dispos.* **29**, 1176-1182.

87 Grant, D. F., Greene, J. F., Pinot, F., Borhan, B., Moghaddam, M. F., Hammock, B. D., McCutchen, B., Ohkawa, H. and Guenthner, T. M. (1996) *Biochem. Pharmacol.* **51**, 503-515.

88 Barnes, H. J., Jenkins, C. M. and Waterman, M. R. (1994) *Arch. Biochem. Biophys.* **315**, 489-494.

89 Romanos, M. (1995) *Curr. Opin. Biotech.* **6**, 527-533.

90 Punt, P. J., van Biezen, N., Conesa, A., Albers, A., Mangnus, J. and van den Hondel, C. (2002) *Trends Biotechnol.* **20**, 200-206.

91 Cregg, J. M., Vedvick, T. S. and Raschke, W. C. (1993) *Biotechnology (NY)* **11**, 905-910.

92 Horiuchi, T., Koga, H. and Sagara, Y. (1993) In: Omura, T., Ishimura, Y. and Fujii-Kuriyama, Y. (Eds), *Cytochrome P450* 2nd ed., pp. 186-205. VCH Publishers Inc., New York.

93 Dong, M. S., Yamazaki, H., Guo, Z. and Guengerich, F. P. (1996) *Arch. Biochem. Biophys.* **327**, 11-19.

94 Jenkins, C. M. and Wateman, M. R. (1994) *J. Biol. Chem.* **269**, 27401-27408.

95 Chun, Y. J., Jeong, T. C., Roh, J. K. and Guengerich, F.P. (1997) *Biochem. Biophys. Res. Commun.* **230**, 211-214.

96 Jenkins, C. M., Pikuleva, I., Kagawa, N. and Waterman, M. R. (1997) *Arch. Biochem. Biophys.* **347**, 93-102.

97 Asseffa, A., Smith, S. J., Nagata, K., Gillette, J., Gelboin, H. V. and Gonzalez, F. J. (1989) *Arch. Biochem. Biophys.* **274**, 481-490.

98 Crespi, C. L. and Miller, V. P. (1999) *Pharmacol. Therapeut.* **84**, 121-131.

99 Romanos, M. A., Scorer, C. A. and Clare J. J. (1992) *Yeast* **8**, 423-488.

100 Oeda, K., Sakaki, T. and Ohkawa, H. (1985) *DNA Cell Biol.* **4,** 167-175.

101 Sakaki, T., Oeda, K., Miyoshi, M. and Ohkawa, H. (1985) *J. Biochem.* **98**, 1667-1675.

102 Cregg, J. M. and Higgins, D. R. (1995) *Can. J. Bot.* **73**, S891-S897.

103 Hollenberg, C. P. and Gellisen, G. (1997) *Curr. Opin. Biotech.* **8**, 554-560.

104 Rosenfeld, S. A. (1999) *Method. Enzymol.* **306**, 154-169.

105 Lin Cereghino, J. and Cregg, J. M. (2000) *FEMS Microbiol. Rev.* **24**, 45-66.

106 Cregg, J. M., Barringer, K. J., Hessler, A. Y. and Madden, K. R. (1985) *Mol. Cell. Biol.* **5**, 3376-3385.

107 Tschopp, J. F., Brust, P. F., Cregg, J. M., Stillman, C., and Gingeras, T. R. (1987). *Nucleic Acids Res.* **15**, 3859-3876.

108 Fisher, C. W., Shet. M. S., Caudle, D. L. and Martin-Wixtrom, C. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10817-10821.

109 Shet, M. S., Fisher, C. W., Holmans, P. L. and Estabrook, R. W. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11748-11752.

110 Fisher, C. W., Shet, M. S. and Estabrook, R. W. (1996) *Method. Enzymol.* **272**, 15-23.

111 Dong, J. and Porter, T. D. (1996) *Arch. Biochem. Biophys.* **327**, 254-259.

112 Shet, M. S., Fisher, C. W. and Estabrook, R. W. (1997) *Arch. Biochem. Biophys.* **339**, 218-225.

113 Voice, M. W., Zhang, Y., Wolf, C. R., Burchell, B. and Friedberg, (1999) *Arch. Biochem. Biophys.* **366**, 116-124.

114 Blake, J. A. R., Pritchard, M., Ding, S., Smith, G. C. M., Burchell, B., Wolf, C. R. and Friedberg, T. (1996) *FEBS Lett.* **397**, 210-214.

115 Gillam, E. M. J., Baba, T., Kim, B.-R., Ohmori, S. and Guengerich, F. P. (1993) *Arch. Biochem. Biophys.* **305**, 123-131.

116 Parikh, A., Gillam, E. M. J. and Guengerich, F. P. (1997) *Nat. Biotechnol.* **15**, 784-788.

117 Pritchard, M. P., Glancey, M. J., Blake, J. A. R., Gilham, D. E., Burchell, B., Wolf, R. C. and Friedberg, T. (1998) *Cancer Res.* **58**, 1833-1838.

118 Paine, M. J. I., Gilham, D., Roberts, G. C. K. and Wolf, C. R. (1996) *Arch. Biochem. Biophys.* **328**, 143-150.

119 Wang, M.-H., Patten, C. J., Yang, G.-Y., Paranawithana, S. R., Tan, Y. and Yang, C. S. (1996) *Arch. Biochem. Biophys.* **334**, 380-388.

120 Urban, P., Cullin, C. and Pompon, D. (1990) *Biochimie* **72**, 463-472.

121 Maughan, J. A., Nugent, J. H. A. and Hallahan, D. L. (1997) *Arch. Biochem. Biophys.* **341**, 104-111.

122 Murakami, H., Yabusaki, Y., Sakaki, T., Shibata, M. and Ohkawa, H. (1987) *DNA Cell Biol.* **6**, 189-197.

123 Sakaki, T., Shibata, M., Yabusaki, Y., Murakami, H. and Ohkawa, H. (1989) *DNA* **8**, 409-418.

124 Murakami, H., Yabusaki, Y., Sakaki, T., Shibata, M. and Ohkawa, H. (1990) *J. Biochem.* **108**, 850-865.

125 Sakaki, T., Kominami, S., Takemori, S., Ohkawa, H., Akiyoshi-Shibata, M. and Yabusaki, Y. (1994) *Biochemistry* **33**, 4933-4939.

126 Sakaki, T., Shibata, M., Yabusaki, Y., Murakami, H. and Ohkawa, H. (1990) *DNA Cell Biol.* **9**, 603-614.

127 Eugster, H.-P. and Sengstag, C. (1993) *Toxicology* **82**, 61-73.

128 Sengstag, C., Eugster, H.-P. and Würgler, F. E. (1994) *Carcinogenesis* **15**, 837-844.

129 Hayashi, K., Sakaki, T., Kominami, S., Inouye, K. and Yabusaki, Y. (2000) *Arch. Biochem. Biophys.* **381**, 164-170.

130 Shibata, M., Sakaki, T., Yabusaki, Y., Murakami, H. and Ohkawa, H. (1990) *DNA Cell Biol.* **9**, 27-36.

131 Urban, P., Truan, G., Bellamine, A., Laine, R., Gautier, J.-C. and Pompon, D. (1994) *Engineered yeast stimulating P450-dependent metabolism: tricks, muths and reality.* In: N. Kingsley (Ed.), Drug metabolism and interactions. Freund Publishing, London, pp.169-200.

Stellenbosch University http://scholar.sun.ac.za

132 Truan, G., Cullin, C,, Reisdorf, P., Urban, P. and Pompon, D. (1993) *Gene* **125**, 49-55.

133 Pompon, D., Louerat, B., Bronine, A. and Urban, P. (1996) *Method. Enzymol.* **272**, 51-63.

134 Gautier, J.-C., Urban, P., Beaune, P. and Pompon, D. (1993) *Eur. J. Biochem.* **211**, 63-72.

135 Urban, P., Truan, G., Gautier, J.-C. and Pompon, D. (1993) *Biochem. Soc. Trans.* **21**, 1028-1034.

136 Ohgiya, S., Goda, T., Hoshino, T., Kamataki, T. and Ishizaki, K. (1997) *Arch. Biochem. Biophys.* **343**, 215-224.

137 Kusano, K., Waterman, M. R., Sakaguchi, M., Omura, T. and Kagawa, N. (1999) *Arch. Biochem. Biophys.* **367**, 129-136.

138 Kagawa, N., Hori, H., Waterman, M. R. and Yoshioka, S. (2004) *Steroids* **69**, 235-243.

139 Inoue, E., Takahashi, Y., Imai, Y. and Kamataki, T. (2000) *Biochem. Biophys. Res. Commun.* **269**, 623-627.

140 Kagawa, N., Cao, Q. and Kusano, K. (2003) *Steroids* **68**, 205-209.

141 Ahn, T., Yang, S. and Yun, C.-H. (2004) *Protein Expres. Purif.* **36**, 48-52.

142 Cole, P. A. (1996) *Structure* **4**, 239-242.

143 Lopez, P. J., Iost, I. and Dreyfus, M. (1994) *Nucleic Acids Res.* **22**, 1186-1193.

144 Brian, W. R., Srivastava, P. K., Umbenhauer, D. R., Lloyd, R. S. and Guengerich, F. P. (1990) *Biochemistry* **28**, 4993-4999.

145 Lopez-Garcia, M. P., Dansette, P. M., Valadon, P., Amar, C., Beaune, P. H., Guengerich, F. P. and Mansuy, D. (1993) *Eur. J. Biochem.* **213**, 223-232.

146 Imaoka, S., Yamada, T., Hiroi, T., Hayashi, K., Sakaki, T., Yabusaki, Y. and Funae, Y. (1996) *Biochem. Pharmacol.* **51**, 1041-1050.

147 Shimizu, T., Sogawa, K., Fujii-Kuriyama, Y., Takahashi, M., Ogoma, Y. and Hatano, M. (1986) *FEBS Lett.* **207**, 217-221.

148 Shimizu, T., Hirano, K., Takahashi, M., Hatano, M. and Fujii-Kuriyama, Y. (1988) *Biochemistry* **27**, 4138-4141.

149 Fujita, V. S., Thiele, D. J. and Coon, M. J. (1990) *DNA Cell Biol.* **9**, 111-118.

150 Pernecky, S. J., Porter, T. D. and Coon, M. J. (1990) *Biochem. Biophys. Res. Commun.* **172**, 1331-1337.

151 George, S. G., Scott, J. and Ellis, S. W. (1995) *J. Mar. Biotechnol.* **3**, 220-223.

152 Ching, M. S., Lennard, M. S., Tucker, G. T., Woods, H. F., Kelly D. E. and Kelly, S. L. (1991) *Biochem. Pharmacol.* **42**, 753-758.

153 Akiyoshi-Shibata, M., Usui, E., Sakaki, T., Yabusaki, Y., Noshiro, M., Okuda, K. and Ohkawa, H. (1991) *FEBS Lett.* **280**, 367-370.

Stellenbosch University http://scholar.sun.ac.za

154 Sakaki, T., Akiyoshi-Shibata, M., Yabusaki, Y. and Ohkawa, H. (1992) *J. Biol. Chem.* **267**, 16497-16502.

155 Ellis, S. W., Ching, M. S., Watson, P. E., Henderson, C. J., Simula, A. P., Lennard, M. S., Tucker, G. T. and Woods, H. F. (1992) *Biochem. Pharmacology* **44**, 617-620.

156 Hashizume, T., Imaoka, S., Hiroi, T., Terauchi, Y., Fujii, T., Miyazaki, H., Kamataki, T. and Funae, Y. (2001) *Biochem. Biophys. Res. Commun.* **280**, 1135-1141.

157 Kingsman, S. M., Kingsman, A. J. and Mellor, J. (1987) *Trends Biotechnol.* **5**, 53-57.

158 Wiseman, A. (1993) *Trends Biotechnol.* **11**, 131-136.

159 Eugster, H.-P., Sengstag, C., Meyer, U. A., Hinnen, A. and Würgler. F. E. (1990) *Biochem. Biophys. Res. Commun.* **172**, 737-744.

160 Moir, D. T. and Davidow, L. S. (1991) *Method. Enzymol.* **194**, 491-519.

161 Clare, J. J., Rayment, F. B., Ballantine, S. P., Sreekrishna, K. and Romanos, M. A. (1991) *Biotechnology (NY)* **9**, 455-460.

162 Thill, G. P., Davis, G. R., Stillman, C., Holtz, G., Brierly, R., Engel, M., Buckholtz, R., Kenney, J., Provow, S., Vedvick, T. and Siegel, R. S. (1990) In: *Proceedings of the Sixth International Symposium on the Genetics of Microorganisms* (Heslot, H., Davies, J., Florent, J., Bobichon, L., Durand, G. and Penasse, L., Eds.), Vol. 2, pp. 477-490. Société Française de Microbiologie, Paris.

163 Cregg, J. M. and Madden, K. R. (1988) *Dev. Ind. Microbiol.* **29**, 33-41.

164 Anderson, M. D., Busk, P. K., Svendsen, I and Møller, B. L. (2000) *J. Biol. Chem.* **275**, 1966-1975.

165 Kolar N. W. (2005) Ph. D dissertation.

166 Manual of Pichia Muti-Copy Expression Kit, Invitrogen, San Diego, CA.

167 Kozak, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2850-2854.

168 Baim, S. B. and Sherman, F. (1988) *Mol. Cell. Biol.* **8**, 1591-1601.

169 Sharp, P. M. and Cowe, E. (1991) *Yeast* **7**, 657-678.

170 Gold, L. and Stormo, G. (1990) *Method. Enzymol.* **185**, 1302-1307.

171 Bennetzen, J. L. and Hall, B. D. (1982) *J. Biol. Chem.* **257**, 3026-3031.

172 Clemens, J. M., Laz, T. and Sherman, F. (1989) In: *Yeast Genetic Engineering*, pp. 65-82. Barr, P. J., Brake, A. J. and Valenzuela, P. (Eds.). Butterworth Publishers, Stoneham, MA.

173 Kozak, M. (1996) *Mamm. Genome* **7**, 563-574.

174 De Smit, M. H. and van Duin, J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 7668-7672.

175 Looman, A. C., Bodlaender, J., de Gruyter, M., Vogelaar, A and van Knippenberg, P. H. (1986) *Nucleic Acids Res.* **14**, 5481-5497.

176 Barnes, H. J., Arlotto, M. P. and Waterman, M. R. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5597-5601.

177 Looman, A. C., Bodlaender, J., Comstock, L. J., Eaton, D., Jhurani, P., de Boer, H. A. and van Knippenberg, P. H. (1987) *EMBO J.* **6**, 2489-2492.

178 Stormo, G. D., Schneider, T. D. and Gold, L. D. (1982) *Nucleic Acids Res.* **10**, 2971-2996.

179 Imai, T., Globerman, H., Gertner, J., Kagawa, N. and Waterman M. R. (1993) *J. Biol. Chem.* **268**, 17317-17375.

180 Palmer, C. N. A., Richardson, T. H., Griffin, K. J., Hsu, M.-H., Muerhoff, A. S., Clark, J. E. and Johnson, E. F. (1993) *BBA-Gene Struct. Expr.* **1172**, 161-166.

181 Nishimoto, M., Clark, J. E. and Masters, B. S. S. (1993) *Biochemistry* **32**, 8863-8870.

182 Andersen, J. F., Utermohlen, J. G. and Feyereisen, R. (1994) *Biochemistry* **33**, 2171-2177.

183 Shimada, T., Wunsch, R. M., Hanna, I. H., Sutter, T. R., Guengerich, F. P. and Gillam, E. M. (1998) *Arch. Biochem. Biophys.* **357**, 111-120.

184 Pernecky, S. J. and Coon, M. J. (1996) *Method. Enzymol.* **272**, 25-34.

185 Li, Y. C. and Chiang, Y. L. (1991) *J. Biol. Chem.* **266**, 19186-19191.

186 Sagara, Y., Barnes, H. J. and Waterman, M. R. (1993) *Arch. Biochem. Biophys.* **304**, 272-278.

187 Clark. B. J. and Waterman, M. R. (1991) *J. Biol. Chem.* **266**, 5898-5904.

188 Larson, J. R., Coon, M. J. and Porter, T. D. (1991) *J. Biol. Chem.* **266**, 7321-7324.

189 Batard, Y., Hehn, A., Nedelkina, S., Schalk, M., Pallett, K., Schaller, H. and Werck-Reichhart, D. (2000) *Arch. Biochem. Biophys.* **379**, 161-169.

190 Hamilton, R., Watanabe, C. K. and de Boer, H. A. (1987) *Nucleic Acids Res.* **15**, 3581-3593.

191 Cullin, C. and Pompon, D. (1988) *Gene* **65**, 203-217.

192 Hayes, C. L., Spink, D. C., Spink, B. C., Cao, J. Q., Walker, N. J. and Sutter, T. R. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 9776-9781.

193 Dobson, M. J., Tuite, M. F., Roberts, N. A., Kingsman, A. J. and Kingsman, S. M. (1982) *Nucleic Acids Res.* **10**, 2625–2637.

194 Eugster, H.-P., Probst, M., Würgler, F. E. and Sengstag, C. (1993) *Drug Metab. Dispos.* **21**, 43-49.

195 Weickert, M. J., Doherty, D. H., Best, E. A. and Olins, P. O. (1996) *Curr. Opin. Biotech.* **7**, 494-499.

Stellenbosch University http://scholar.sun.ac.za

196 Richardson, T. H., Jung, F., Griffin, K. J., , M., Raucy, J. L., Kemper, B., Bornheim, L. M., Hassett, C., Omiecinski, C. J. and Johnson, E. F. (1995) *Arch. Biochem. Biophys.* **323**, 87-96.

197 Jansson, I., Stoilov, I., Sarfarazi, M. and Schenkman, J. B. (2000) *Toxicology* **144**, 211-219.

198 Gillam, E. M. J., Guo, Z., Martin, M. V., Jenkins, C. M. and Guengerich, F. P. (1995) *Arch. Biochem. Biophys.* **319**, 540-550.

199 Kempf, A. C., Zanger, U. M. and Meyer, U. A. (1995) *Arch. Biochem. Biophys.* **321**, 277-288.

200 Ness, F., Achstetter, T., Duport, C., Karst, F., Spagnoli, R. and Degryse, E. (1998) *J. Bacteriol.* **180**, 1913-1919.

201 Guengerich, F. P., Brian, W. R., Sari, M.-A. and Ross, J. T. (1991) *Method. Enzymol.* **206**, 130-145.

202 Nishihara, H., Okamura, T., Schmid, R. D., Hauck, A. and Reuß, M. (1997) *J. Biotechnol.* **56**, 57-61.

203 Siegel, R. S. and Brierley, R. A. (1989) *Biotechnol. Bioeng.* **34**, 403-404.

204 Brierley, R. A., Bussineau, C., Kosson, R., Melton, A. and Siegel, R. S. (1990) *Biotechnol. Bioeng.* **589**, 350-362.

205 Stratton, J., Chiruvolu, V. and Meagher, M. (1998) In: *Pichia protocols* (D. R. Higgins and J. M. Cregg, eds.) Humana Press, Inc., Totawa, NJ.

206 Gierasch, L. M. (1989) *Biochemistry* **28**, 923-930.

207 Pritchard, M. P., Ossetian, R., Li, D. N., Henderson, J., Burchell, B., Wolf, C. R. and Friedberg, T. (1997) *Arch. Biochem. Biophys.* **345**, 342-354.

208 Sreekrishna, K., Brankamp, R. G., Kropp, K. E., Blankenship, D. T., Tsay, J.-T., Smith, P. L., Wierschke, J. D., Subramaniam, A. and Birkenberger, L. A. (1997) *Gene* **190**, 55-62.

209 Lin Cereghino, G. P., Lin Cereghino, J., Ilgen, C. and Cregg, J. M. (2002) *Curr. Opin. Biotech.* **13**, 329-332.

210 Poulos, T. L., Finzel, B. C. and Howard, A. J. (1986) *Biochemistry* **25**, 5314-5322.

211 Boddupalli, S. S., Hasemann, C. A., Ravichandran, K. G., Lu, J.-Y., Goldsmith, E. J., Deisenhofer, J. and Peterson, J. A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 5567-5571.

212 Ravichandran, K. G., Boddupalli, S. S., Hasemann, C. A., Peterson, J. A., and Deisenhofer, J. (1993) *Science* **261**, 731-736.

213 Hasemann, C. A., Ravichandran, K. G., Peterson, J. A. and Deisenhofer, J. (1994) *J. Mol. Biol.* **236**, 1169-1185.

214 Cupp-Vickery, J. R. and Poulos, T. L. (1995) *Nat. Struct. Biol.* **2**, 144-153.

Stellenbosch University http://scholar.sun.ac.za

215 Ito, S., Matsuoka, T., Watanabe, I., Kagasaki, T., Seriwaza, T. and Hata T. (1999) *Acta Crystallogr. D.* **55**, 1209-1211.

216 Park, S. Y., Yamane, K., Adachi, S., Shiro, Y., Weiss, K. E. and Sligar, S. G. (2000) *Acta Crystallogr. D* **56** 1173-1175.

217 Meharenna, Y. T., Li, H., Hawkes, D. B., Pearson, A. G., De Voss, J. and Poulos, T. L. (2004) *Biochemistry* **43**, 9487-9494.

218 Park, S. M., Shimizu, H., Adachi, S., Nakagawa, A., Tanaka, I., Nakahara, K., Shoun, H., Obayashi, E., Nakamura, H., Iizuka, T. and Shiro, Y. (1997) *Nat. Struct. Biol.* **4**, 827-832.

219 Pernecky, S. J., Larson, J. R., Philpot, R. M. and Coon, M. J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2651-2655.

220 von Wachenfeldt, C., Richardson, T. H., Cosme, J. and Johnson, E. F. (1997) *Arch. Biochem. Biophys.* **339**, 107-114.

221 Scheller, U., Kraft, R. Schröder, K. L. and Schunck, W. H. (1994) *J. Biol. Chem.* **269**, 12779-12783.

222 Sueyoshi, T., Park, L. J., Moore, R., Juvonen, R. O. and Negishi, M. (1995) *Arch. Biochem. Biophys.* **322**, 265-271.

223 Williams, P. A., Cosme, J., Sridhar, V., Johnson, E. F. and McRee, D. E. (2000) *Mol. Cell* **5**, 121-131.

224 Wester, M. R., Stout, C. D. and Johnson, E. F. (2002) *Method. Enzymol.* **357**, 73-79.

225 Scott, E. E., Spatzenegger, M. and Halpert, J. R. (2001) *Arch. Biochem. Biophys.* **395**, 57-68.

226 Scott, E. E., He, Y. A., Wester, M. R., White, M. A., Chin, C. C., Halpert, J. R., Johnson, E. F. and Stout, C. D. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 13196-13201.

227 Schoch, G. A., Yano, J. K., Wester, M. R., Griffin, K. J., Stout, C. D. and Johnson, E. F. (2004) *J. Biol. Chem.* **279**, 9497-9503.

228 Williams, P. A., Cosme, J., Ward, A., Angove, H. C., Vinković, D. M. and Jhoti, H. (2003) *Nature* **424**, 464-468.

229 Cosme, J. and Johnson, E. F. (2000) *J. Biol. Chem.* **275**, 2545-2553.

230 Schenkman, J. B. and Jansson, I. (2003) *Pharmacol. Therapeut.* **97**, 139-152.

231 Yamazaki, H., Gillam, E. M., Dong, M. S., Johnson, W. W., Guengerich, F. P. and Shimada, T. (1997) *Arch. Biochem. Biophys.* **342**, 329-337.

232 Yamazaki, H., Ueng, Y.-F., Shimada, T. and Guengerich, F. P. (1995) *Biochemistry* **34**, 8380-8389.

233 Loughran, P. A., Roman, L. J., Miller, R. T. and Masters, B. S. S. (2001) *Arch. Biochem. Biophys.* **385**, 311-321.

234 Guryev, O. L., Gilep, A. A., Usanov, S. A. and Estabrook, R. W. (2001) *Biochemistry* **40**, 5018-5031.

235 Mason, J. I. (1993) *Biochem. Soc. T.* **21**, 1057-1060.

236 Halpert, J. R. (1995) *Annu. Rev. Pharmacol.* **35**, 29-53.

237 Guengerich, F. P. (2002) *Nat. Rev. Drug Discov.* **1**, 359-366.

238 Van Wauwe, J. P. and Janssen, P. A. (1989) *J. Med. Chem.* **32**, 2231-2239.

239 Hartmann, R. W. (1994) *Eur. J. Pharm. Sci.* **2**, 15-16.

240 Brodie, A. M. H. and Njar, V. C. O. (2000) *Steroids* **65**, 171-179.

241 Santen, R. J. (2003) *Steroids* **68**, 559-567.

242 Bulun, S. E., Zeitoun, K. M., Takayama, K., Simpson, E. and Sasano, H. (2000) *Trends Endocrin. Met.* **11**, 22-27.

243 Séralini, G.-E. and Moslemi, S. (2001) *Mol. Cell. Endocrinol.* **178**, 117-131.

244 Lombardi, P. (2002) *Biochim. Biophys. Acta* **1587**, 326-337.

245 Cavalli, A. and Recanatini, M. (2002) *J. Med. Chem.* **45**, 251-254.

246 Brodie, A. (2002) *Trends Endocrin. Met.* **13**, 61-65.

247 Gelboin, H. V. (1993) *Pharmacol. Rev.* **45**, 413-453.

248 Gelboin, H. V., Krausz, K. W., Gonzalez, F. J. and Yang, T. J. (1999) Trends Pharmacol. Sci. **20**, 432-438.

249 Simpson, E. R., Mahendroo, M. S., Means, G. D., Kilgore, M. W., Hinshelwood, M. M., Graham-Lorence, S., Amarneh, B., Ito, Y., Fisher, C. R., Michael, M. D., Mendelson, C. R. and Bulun, S. E. (1994) *Endocr. Rev.* **15**, 342-355.

250 Santen R. J. and Harvey, A. (1999) *Endocr.-Relat. Cancer* **6**, 72-92.

251 Subramanian, A. (2002) *Mol. Biotechnol.* **20**, 41-47.

252 Mendelson, C. R., Wright, E. E., Evans, C. T., Porter, J. C. and Simpson, E. R. (1985) *Arch. Biochem. Biophys*. **243**, 480-491.

253 Yoshida, N. and Osawa, Y. (1991) *Biochemistry* **30**, 3003-3010.

254 Burgess, R. R. and Thompson, N. E. (2002) *Curr. Opin. Biotech.* **13**, 304-308.

255 Kovaleva, I. E., Krynetskii, E. Y. and Luzikov, V. N. (1996) *Biochem. Biophys. Res. Commun.* **221**, 129-132.

256 Powell, K. A., Ramer, S. W., del Cardayre, B., Stemmer, W. P. C., Tobin, M. B., Longchamp, P. F. and Huisman, G. W. (2001) *Angew. Chem. Int. Ed.* **40**, 3948-3959.

257 van Beilen, J. B. and Li, Z. (2002) *Curr. Opin. Biotech.* **13**, 338-344.

258 Szczebara, F. M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Duport, C., Blanchard, S., Groisillier, A., Testet, E., Costaglioli, P., Cauet, G., Degryse, E., Balbuena, D., Winter, J., Achstetter, T., Spagnoli, R., Pompon, D. and Dumas. D. (2003) *Nat. Biotechnol.* **21**, 143-149.

Stellenbosch University http://scholar.sun.ac.za

259 Friedberg, T. and Wolf, C. R. (1996) *Adv. Drug Deliver. Rev.* **22**, 187-213.

260 Belloc, C., Baird, S., Cosme, J., Lecoeur, S., Gautier, J.-C., Challine, D., de Waziers, I., Flinois, J.-P. and Beaune, P. H. (1996) *Toxicology* **106**, 207-219.

261 Sudbery, P. E. (1996) *Curr. Opin. Biotech.* **7**, 517-524.

262 Guo, Z., Gillam, E. M. J., Ohmori, S., Tukey, R. H. and Guengerich, F. P. (1994) *Arch. Biochem. Biophys.* **312**, 436-446.

263 Chun, Y.-J., Shimada, T. and Guengerich, F. P. (1996) *Arch. Biochem. Biophys.* **330**, 48-58.

264 Iwata, H., Fujita, K., Kushida, H., Suzuki, A., Konno, Y., Nakamura, K., Fujino, A. and Kamataki, T. (1998) *Biochem. Pharmacol.* **55**, 1315-1325.

265 Fujita, K., Nakayama, K., Yamazaki, Y., Tsuruma, K., Yamada, M., Nohmi, T. and Kamataki, T. (2001) *Environ. Mol. Mutagen.* **38**, 329-338.

266 Yamazaki, Y., Fujita, K., Nakayama, K., Suzuki, A., Nakamura, K., Yamazaki, H. and Kamataki, T. (2004) *Mutat. Res.* **562**, 151-162.

267 Fisher, C. W., Caudle, D. L., Martin-Wixtrom, C., Quattrochi, L. C., Tukey, R. H., Waterman, M. R. and Estabrook, R. W. (1992) *FASEB J.* **6**, 759-764.

268 Sandhu, P., Guo, Z., Baba, T., Martin, M. V., Tukey, R. H. and Guengerich, F. P. (1994) *Arch. Biochem. Biophys.* **309**, 168-177.

269 Kranendonk, M., Mesquita, P., Laires, A., Vermeulen, N. P. E. and Rueff, J. (1998) *Mutagenesis* **13**, 263-269.

270 Suzuki, A., Kushida, H., Iwata, H., Watanabe, M., Nohmi, T., Fujita, K., Gonzalez, F. J. and Kamataki, T. (1998) *Cancer Res.* **58**, 1833-1838.

271 Soucek, P. (1999) *Arch. Biochem. Biophys.* **370**, 190-200.

272 Hanna, I. H., Reed, J. R., Guengerich, F. P. and Hollenberg, P. F. (2000) *Arch. Biochem. Biophys.* **376**, 206-216.

273 Ariyoshi, N., Miyazaki, M., Toide, K., Sawamura, Y. and Kamataki, T. (2001) *Biochem. Biophys. Res. Commun.* **281**, 1256-1269.

274 Chen, C.-D., Doray, B. and Kemper, B. (1998) *Arch. Biochem. Biophys.* **350**, 233-238.

275 Sandhu, P., Baba, T. and Guengerich, F. P. (1993) *Arch. Biochem. Biophys.* **306**, 443-450.

276 Deeni, Y. Y., Paine, M. J. I., Ayrton, A. D., Clarke, S. E., Chenery, R. and Wolf, R. (2001) *Arch. Biochem. Biophys.* **396**, 16-24.

277 Winters, D. K. and Cederbaum, A. I. (1992) *Biochim. Biophys. Acta* **1156**, 43-49.

278 Gillam, E. M. J., Guo, Z. and Guengerich, F. P. (1994) *Arch. Biochem. Biophys.* **312**, 59-66.

279 Gillam, E. M. J., Guo, Z., Ueng, Y. F., Yamazaki, H., Cock, I., Reilly, P. E. B., Hooper, W. D. and Guengerich, F. P. (1995) *Arch. Biochem. Biophys.* **317**, 374-384.

280 Domanski, T. L., Finta, C., Halpert, J. R. and Zaphiropoulos, P. G. (2001) *Mol. Pharmacol.* **59**, 386-392.

281 Zhang, F., Zhou, D., Kao, Y.-C. and Chen, S. (2002) *Biochem. Pharmacol.* **64**, 1317-1324.

282 Hu, M.-C. and Chung, B.-C. (1990) *Mol. Endocrinol.* **4**, 893-898.

283 Mast, N., Norcross, R., Andersson, U., Shou, M., Nakayama, K., Bjorkhem, I. and Pikuleva, I. A. (2003) *Biochemistry* **42**, 14284-14292.

284 Mast, N., Andersson, U., Nakayama, K., Bjorkhem, I. And Pikuleva, I. A. (2004) *Arch. Biochem. Biophys.* **428**, 99-108.

285 Holmans, P. L., Shet, M. S., Martin-Wixtrom, C. A., Fisher, C. W. and Estabrook, R. W. (1994) *Arch. Biochem. Biophys.* **312**, 554-565.

286 Woods, S. T., Sadleir, J., Downs, T., Triantopoulos, T., Headlam, M. J. and Tuckey, R. C. (1998) *Arch. Biochem. Biophys.* **353**, 109-115.

287 John, G. H., Hasler, J. A., He, Y.-a. and Halpert, J. R. (1994) *Arch. Biochem. Biophys.* **314**, 367-375.

288 Licad-Coles, E, He, K., Yin, H. and Correia, M. A. (1997) *Arch. Biochem. Biophys.* **338**, 35-42.

289 LeBrun, L. A., Xu, F., Kroetz, D. L. and Ortiz de Montellano, P. R. (2002) *Biochemistry* **41**, 5931-5937.

290 Gilep, A. A., Guryev, O. L., Usanov, S. A. and Estabrook, R. W. (2001) *Arch. Biochem. Biophys.* **390**, 222-234.

291 Shet, M. S., Fisher, C. W., Arlotto, M. P., Shackleton, C. H., Holmans, P. L., Martin-Wixtrom, C. A., Saeki, Y. and Estabrook, R. W. (1994) *Arch. Biochem. Biophys.* **311**, 402-417.

292 Savas, Ü., Carstens, C.-P. and Jefcoate, C. R. (1997) *Arch. Biochem. Biophys.* **347**, 181-192.

293 Owaki, A., Takamasa, A., Yamazaki, T and Kominami, S. (2002) *J. Steroid Biochem.* **81**, 255-262.

294 Yamazaki, T., Ohno, T., Sakaki, T., Akiyoshi-Shibata, M., Yabusaki, Y., Imai, T. and Kominami, S. (1998) *Biochemistry* **37**, 2800-2806.

295 Szklarz, G. D., He, Y. Q., Kedzie, K. M., Halpert, J. R. and Burnett, V. L. (1996) *Arch. Biochem. Biophys.* **327**, 308-318.

296 Richardson, T. H, Hsu, M.-H., Kronbach, T., Barnes, H. J., Chan, G., Waterman, M. R., Kemper, B. and Johnson, E. F. (1993) *Arch. Biochem. Biophys.* **300**, 510-516.

297 Tom, M., Myers, C. R. and Waterman M. R. (2002) *Aquat. Toxicol.* **59**, 101-114.

298 Haudenschild, C., Schalk, M., Karp, F. and Croteau, R. (2000) *Arch. Biochem. Biophys.* **379**, 127-136.

299 Hotze, M., Schröder, G. and Schröder G. (1995) *FEBS Lett.* **374**, 345-350.

Stellenbosch University http://scholar.sun.ac.za

300 Halkier, B. A., Nielson, H. L., Koch, B. and Møller, B. L. (1995) *Arch. Biochem. Biophys.* **322**, 369-377.

301 Unger, B. P., Gunsalus, I. C. and Sligar, S. G. (1986) *J. Biol. Chem.* **261**, 1158-1163.

302 Hawkes, D. B., Adams, G. W., Burlingames, A. L., Ortiz de Montellano, P. R. and De Voss, J. J. (2002) *J. Biol. Chem.* **277**, 27725-27732.

303 Green, A. J., Rivers, S. L., Cheesman, M., Reid, G. A., Quaroni, L. G., Macdonald, I. D. G., Chapman, S. K. and Munro, A. W. (2001) *J. Biol. Inorg. Chem.* **6**, 523-533.

304 Buters, J. T., Shou, M., Hardwick, J. P., Korzekwa, K. R. and Gonzalez, F. J. (1996) *Drug Metab. Dispos.* **23**, 696-701.

305 Zeldin, D. C., DuBois, R. N., Falck, J. R. and Capdevila, J. H. (1995) *Arch. Biochem. Biophys.* **232**, 76-86.

306 Grogan, J., Shou, M., Andrusiak, E. A., Tamura, S., Buters, J. T. M., Gonzalez, F. J. and Korzekwa, K. R. (1995) *Biochem. Pharmacol.* **50**, 1509-1515.

307 Biagini, C. and Celier, C. (1996) *Arch. Biochem. Biophys.* **326**, 298-305.

308 Patten, C. J. and Koch, P. (1995) *Arch. Biochem. Biophys.* **317**, 504-513.

309 Chen, W., Peter, R. M., McArdle, S., Thummel, K. E., Sigle, R. O. and Nelson, S. D. (1996) *Arch. Biochem. Biophys.* **335**, 123-130.

310 Wu, S., Moomaw, C. R., Tomer, K. B., Falck, J. R. and Zeldin, D. C. (1996) *J. Biol. Chem.* **271**, 3460-3468.

311 Buters, J. T. M., Korzekwa, K. R., Kunze, K. L., Omata, Y., Hardwick, J. P. and Gonzalez, F. P. (1994) *Drug Metab. Dispos.* **22**, 688-692.

312 Lee, C. A., Kadwell, S. H., Kost, T. A. and Serabjit-Singh, C. J. (1995) *Arch. Biochem. Biophys.* **319**, 157-167.

313 Sakuma, T., Kitamura, R., Yokoi, T. and Kamataki, T. (1995) *Biochem. Mol. Biol. Int.* **35**, 447-455.

314 Imaoka, S., Ogawa, H., Kimura, S. and Gonzalez, F. J. (1993) *DNA Cell Biol.* **12**, 893-899.

315 Shimozawa, O., Sakaguchi, M., Ogawa, H., Harada, N., Mihara, K. and Omura, T. (1993) *J. Biol. Chem.* **268**, 21399-21402.

316 Lähde, M., Raunio, H., Pelkonen, O., Karp, M. and Oker-Blom, C. (1993) *Biochem. Biophys. Res. Commun.* **197**, 1511-1517.

317 Sigle, R. O., Titus, M. A., Harada, N. and Nelson, S. D. (1994) *Biochem. Biophys. Res. Commun.* **201**, 694-700.

318 Amarneh, B. and Simpson, E. R. (1995) *Mol. Cell. Endocrinol.* **109**, R1-R5.

319 Gartner, C. A., Thompson, S. J., Rettie, A. E. and Nelson, S. D. (2001) *Protein Expres. Purif.* **22**, 443-454.

Stellenbosch University http://scholar.sun.ac.za

320 Zhang, Q.-Y., Ding, X. and Kaminsky, L. S. (1997) *Arch. Biochem. Biophys.* **340**, 270-278.

321 Wittekindt, N. E., Würgler, F. E. and Sengstag, C. (1995) *DNA Cell Biol.* **14**, 273-283.

322 Yasumori, T., Murayama, N., Yamazoe, Y., Nogi, T., Fukasawa, T, and Kato, R. (1989) *Mol. Pharmacol.* **35**, 443-449.

323 Goldstein, J. A., Faletto, M. B., Romkes-Sparks, M., Sullivan, T. and Kitareewan, S. (1994) *Biochemistry* **33**, 1743-1752.

324 Krynetsky, E.Y., Drutsa, V.L., Kovaleva, I.E., Luzikov, V.N. and Uvarov, V.Y. (1993) *FEBS Lett.* **336**, 87-89.

325 Brian, W. R., Sari, M.-A., Iwasaki, M., Shimada, T., Kaminsky, L. S. and Guengerich, F. P. (1990) *Biochemistry* **29**, 11280-11292.

326 Renaud, J.-P., Cullin, C., Pompon, D., Beaune, P. and Mansuy, D. (1990) *Eur. J. Biochem.* **194**, 889-896.

327 Peyronneau, M.-A., Renaud, J.-P., Jaouen, M., Urban, P. and Cullin, C. (1993) *Eur. J. Biochem.* **218**, 355-361.

328 Bylund, J., Bylund, M. and Oliw, E. H. (2001) *Biochem. Biophys. Res. Commun.* **280**, 892-897.

329 Swart, A. C., Swart, P., Roux, S. P.,van der Merwe, K. J., Pretorius, I. S. and Steyn, A. J. C. (1995) *Endocr. Res.* **21**, 289-295.

330 Pompon, D., Liu, R. Y.-K., Besman, M. J., Wang, P.-L., Shively, J. E. and Chen, S. (1989) *Mol. Endocrinol.* **3**. 1477-1487.

331 Wu, D.-A., Hu, M.-C. and Chung, B.-C. (1991) *DNA Cell Biol.* **10**, 201-209.

332 Lundell, K. (2002) *Biochem. J.* **363**, 297-303.

333 Nagata, K., Matsunaga, T., Gillette, J., Gelboin, H. V. and Gonzalez, F. J (1987) *J. Biol. Chem.* **262**, 2787-2793.

334 Hayashi, S., Morohashi, K., Yoshioka, H., Okuda, K. and Omura, T. (1988) *J. Biochem.* **103**, 858-862.

335 Kimura, S., Smith, H. H., Hankinson, O. and Nebert, D. W. (1987) *EMBO J.* **6**, 1929-1933.

336 Pompon, D. and Nicolas, A. (1989) *Gene* **83**, 15-14.

337 Pompon, D. (1988) *Eur. J. Biochem.* **177**, 285-293.

338 Kedzie, K. M., Philpot, R. M. and Halpert, J. R. (1991) *Arch. Biochem. Biophys.* **291**, 176-186.

339 Imai, Y. and Nakamura, M. (1988) *FEBS Lett.* **234**, 313-315.

340 Yokotani, N., Bernhardt, R., Sogawa, K., Kusunose, E., Gotoh, O., Kusunose, M. and Fujii-Kuriyama, Y. (1989) *J. Biol. Chem.* **264**, 21665-21669.

Stellenbosch University http://scholar.sun.ac.za

341 Wiedmann, B., Silver, P., Schunck, W.-H. and Wiedmann, M. (1993) *Biochim. Biophys. Acta* **1153**, 267-276.

342 Scheller, U., Zimmer, T., Kärgel, E and Schunck, W.-H. (1996) *Arch. Biochem Biophys.* **328**, 245-254.

343 Siminszky, B., Corbin, F. T., Ward, E. R., Fleischmann, T. J. and Dewey, R. E. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 1750-1755.

344 Urban, P., Werck-Reichhart, D., Teutsch, H. G., Durst, F., Regnier, S., Kazmaier, M. and Pompon, D. (1994) *Eur. J. Biochem.* **222**, 843-850.

345 Urban, P., Mignotte, C., Kazmaier, M, Delorme, F. and Pompon, D. (1997) *J. Biol. Chem.* **272**, 19176-19186.

346 Schoendorf, A., Rithner, C. D., Williams, R. M. and Croteau, R. B. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 1501-1506.

347 Smith, F. F. and Scott, J. G. (1997) *Insect Biochem. Molec.* **27**, 999-1006.

348 Boyle, S. M., Popp, M. P., Smith, W. C., Greenberg, R. M. and James, M. O. (1998) *Mar. Environ. Res.* **46**, 25-28.

349 Trant, J. M. (1996) *Arch. Biochem. Biophys.* **326**, 8-14.

350 Cleves, A. E. and Bankaitis V. A. (1992) In: *Adv. Microb. Physiol.*, **33**, 73-144.

351 Schekman, R. and Novick, P. (1982) In : *The Molecular biology of the Yeast Saccharomyces*, pp. 361-393. Strathern, J. N., Jones, E. W. and Broach, J. R. (Eds.). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

352 Sleep, D., Belfield, G. P. and Goodey, A. R. (1990) *Biotechnology (NY)***8**, 42-46.

353 Smith, R. A., Duncan, M. J. and Moir, D. T. (1985) *Science* **229**, 1219- 1224.

354 Romanos, M.A, Scorer, C.A and Clare, J.J. (1995) *DNA Cloning 2* (D. M. Glover and B. D. Hames, ed.) 2nd Ed, Oxford University Press Inc., New York.

355 Walter, P. and Lingappa, V. R. (1986) *Annu. Rev. Cell Bio.* **2**, 499-516.

356 Rothman, J. E. and Orci, L. (1990) *FASEB J.* **4**, 1460-1468.

357 Rothman, J. E. (1994) *Nature* **372**, 55-63.

358 Kaiser, C. A., Gimeno, R. E. and Shaywitz, D. A. (1997) In : *The molecular and cellular biology of the yeast Saccharomyces*, pp. 91-228. Pringle, J. R., Jones, E.W. and Broach, J.R. (Eds.). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

359 Sakaguchi, M. (1997) *Curr. Opin. Biotech.* **8**, 595-601.

360 Hitzeman, R. A., Chen, C. Y., Dowbenko, D. J., Renz, M. E., Liu, C., Pai, R., Simpson, N. J., Kohr, W. J., Singh, A., Chisholm, V., Hamilton, R. and Chang, C. N. (1990) *Method. Enzymol.* **185**, 421-440.

361 Kaiser, C. A., Preuss, D., Grisafi, P. and Botstein, D. (1987) *Science* **235**, 312-317.

362 Sakai, A., Shimiza, Y. and Hishinama, F. (1988) *Genetics* **119**, 499-506.

363 Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlein, U. A., Masiarz, F. R., Mullenbach, G. T., Urdea, M. S., Valenzuela, P. and Barr, P. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4642-4646.

364 Zsebo, K. M., Lu, H. S., Fieschko, J. C., Goldstein, L., Davis, J., Duker, K., Suggs, S. V., Lai, P.-H. and Bitter, G. A. (1986) *J. Biol. Chem.* **261**, 5858-5865.

365 Brake, A. J. (1990) *Method. Enzymol.* **185**, 408-421.

366 Green, R., Kramer, R. A. and Shields, D. (1989) *J. Biol. Chem.* **264**, 2963-2968.

367 von Heijne, G. (1990) *J. Membrane Biol.* **115**, 195-201.

368 Martoglio, B. and Dobberstein, B. (1998) *Trends Cell Biol.* **8**, 410- 415.

369 Ng, T. W., Brown, J. D. and Walter, P. (1996) *J. Cell Biol.* **134**, 269-278.

370 Siegal, V. and Walter, P. (1985) *J. Cell. Biol.* **100**, 1913-1921.

371 Wiedmann, B., Sakai, H., Davis, T. A. and Wiedmann, M. (1994) *Nature* **370**, 434-440.

372 Powers, T. and Walter, P. (1996) *Curr. Biol.* **6**, 331-338.

373 Crowley, K. S., Liao, S., Worrell, V. E., Reinhart, G. D. and Johnson, A. E. (1994) *Cell* **78**, 461-471.

374 Gething, M. J. and Sambrook, J. (1992) *Nature* **355**, 33-45.

375 Sidrauski, C., Chapman, R. and Walter, P. (1998) *Trends Cell Biol.* **8**, 245-249.

376 Innis, M. A. (1989) In: P. J Barr, A. J. Brake and P. Valenzuela, P. (Eds), *Yeast Genetic Engineering*, pp. 233-246. Butterworths, Boston.

377 Balou, C. E. (1990) *Method. Enzymol.* **185**, 440-470.

378 Caplan, S., Green, R., Rocco, J. and Kurjan, J. (1991) *J. Bacteriol.* **173**, 627-635.

379 Beckers, C. J. M., Plutner, H., Davidson, H. W. and Balch, W. E. (1990) *J. Biol. Chem.* **265**, 18298-18310.

380 Gorelick, F. S. and Shugrue, C. (2001) *Mol. Cell. Endrocrinol.* **177**, 13-18.

381 Baker, D., Hicke, L., Rexach, M., Schleyer, M. and Schekman, R. (1988) *Cell* **54**, 335-344.

382 Griffiths, G. and Simons, K. (1986) *Science* **234**, 438-443.

383 Franzusoff, A. Rothblatt, J. and Schekman, R. (1991) *Method. Enzymol.* **194**, 662-674.

384 Jungmann, J. and Munro, S. (1998) *EMBO J.* **17**, 423-434.

385 Kornfeld, R. and Kornfeld, S. (1985) *Annu. Rev. Biochem.* **54**, 631-664.

386 Munro, S. (2001) *FEBS Lett.* **498**, 223-227.

387 Rothman, J. E. (1994) *Nature* **372**, 55-63.

388 Salminen, A. and Novick, P. J. (1987) *Cell* **49**, 527-538.

389 Goud, B., Salminen, A., Walworth, N. C. and Novick, P. J. (1988) *Cell* **53**, 753-768.

390 Walworth, N. C., Goud, B., Kabcenell, A. K. and Novick, P. J. (1989) *EMBO J.* **8**, 1685-1693.

391 Ramírez, M., Munoz, M. D. and Larriba, G. (1989) *Biochim. Biophys. Acta* **990**, 206-210.

392 Tanner, W. and Lehle, L. (1987) *Biochim. Biophys. Acta* **906**, 81-99.

393 De Noble, J. G., Klis, F. M., Munnik, T., Priem, J. and van den Ende, H. (1990) *Yeast* **6**, 483-490.

394 De Nobel, J. G., Dijkers, C., Hooijberg, E. and Klis, F. M. (1989) *J. Gen. Microbiol.* **135**, 2077-2084.

395 De Nobel, J. G., Klis, F. M., Priem, J., Munnik, T. and van den Ende, H. (1990) *Yeast* **6**, 491-499.

396 Orlean, P. (1997) In : *The molecular and cellular biology of the yeast Saccharomyces*, pp. 91-228. Pringle, J. R., Jones, E.W. and Broach, J.R. (Eds.). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Stellenbosch University http://scholar.sun.ac.za

PART II

CLONING AND EXTRACELLULAR EXPRESSION OF RECOMBINANT HUMAN CYTOCHROME P45017α (17α-HYDROXYLASE/17,20-LYASE) IN *SACCHAROMYCES CEREVISIAE*

Stellenbosch University http://scholar.sun.ac.za

INTRODUCTION **95**

CHAPTER 1 : CLONING AND CONSTRUCTION OF THE FULL-LENGTH HUMAN CYP17 INTO YEPFLAG EXPRESSION VECTOR AND TRANSFORMATION IN *S. CEREVISIAE* **96**

1.1 Materials and Methods **96**

1.1.1 Cloning and construction of the pFLAG-CYP17 expression vector 96

1.1.1.1 Polymerase chain reaction (PCR) 96

1.1.1.2 Restriction digestion 96

1.1.1.3 Ligation 96

1.1.1.4 Transformation 96

1.1.2 Analyses of *E. coli* transformants 96

1.1.2.1 PCR amplification of *E. coli* transformants 96

1.1.2.2 DNA sequencing strategy of *E. coli* transformants 96

1.1.2.3 Southern hybridisation of CYP17 DNA 96

1.1.3 *S. cerevisiae* transformation and screening of transformants 96

1.2 Results **96**

1.2.1 Confirmation of the integrity of the YEpFLAG-1 plasmid 96

1.2.2 Analyses of *E. coli* transformants 96

1.2.3 Analyses of *S. cerevisiae* transformants 96

CHAPTER 2 : CULTURE OF *S. CEREVISIAE* EXPRESSING P45017α AND OPTIMISATION OF PROTEIN EXPRESSION **96**

2.1 Materials and Methods **96**

2.1.1 Growth and optimisation of *S. cerevisiae* P45017α expression 96

2.1.2 Evaluation of P45017α expression at transcriptional level 96

2.1.2.1 RNA isolation 96

2.1.2.2 Northern hybridisation of CYP17 mRNA 96

2.1.2.3 RNA dot blot 96

2.1.3 P45017α protein expression 96

Stellenbosch University http://scholar.sun.ac.za

2.1.3.1 Extracellular protein expression 96

2.1.3.2 Preparation of the intracellular protein fraction 96

2.1.3.3 Preparation of the periplasmic and intracellular protein fractions 96

2.1.4 SDS-PAGE and Western blot analyses 96

2.2 Results 96

2.2.1 Growth of *S. cerevisiae* 96

2.2.2 Transcription levels of CYP17 in yeast during induction phase 96

2.2.3 Intra- and extracellular expression of P45017α 96

CHAPTER 3 : DISCUSSION 96

REFERENCES 96

Stellenbosch University http://scholar.sun.ac.za

INTRODUCTION

Due to the low level of expression and instability outside the lipid environment of the membrane, human P45017α has not been previously purified from human tissue. The limited availability of human tissue and the genetic variability between human samples has also complicated matters in this regard. To overcome these problems researchers have been using several heterologous expression systems, such as COS cells (1, 2, 3, 4, 5, 6, 7), baculovirus (8), yeast (9, 10, 11, 12) and *E. coli* (13, 14, 15, 16), to produce an active human P45017α protein (17). Of these expression systems only the bacterial system was capable of producing high yields of P45017α protein (18). However, in *E. coli* expression systems most P450 cDNA, including the human CYP17 cDNA, sequence had to be modified at the amino terminus to achieve adequate levels of expression (13). The most frequently used modification was the replacement of the native P450 amino-terminus with the sequence (MALLLAVF) derived from bovine P45017α coding sequence (13), which created a preferred translational initiation context in *E. coli* and minimised the potential for unwanted secondary structure formation. Although, these small changes in the membrane binding domain were shown to have no significant effect on the catalytic properties of cytochromes P450 (19), the alterations might influence the secondary refolding of the membrane protein and therefore crystallisation might not reflect the true secondary and tertiary structures when compared to an unmodified protein. Secondly, the antibodies raised against an altered or truncated form of human P45017α might be slightly different since the immunogenicity will differ.

The aim of this study was to express an unmodified human P45017α protein, fused at the amino-terminus to a yeast secretory signal peptide, thereby producing high yields of a pure full-length human P45017α protein. The initial aim was to obtain sufficient protein for antibody production alone and preservation of activity was therefore not a priority. The use of signal peptides to express the human cytochromes P450 proteins (P450 2A6, P450 2E1 and P450 3A4) in *E. coli,* without changes to the amino-terminus of the P450 coding sequence, was demonstrated by Pritchard *et al*. (20). Although this procedure resulted in a three times more spectrally active P450 enzymes in bacterial

Stellenbosch University http://scholar.sun.ac.za

cells, the membrane P450 content was not higher compared to traditional amino-terminal modified bovine P45017α P450 2E1 and P450 2A6 (21, 22). This system, in addition, did not remove the signal peptide successfully during its translocation across the bacterial inner membrane, producing an unprocessed protein.

It was envisaged that the yeast expression system initially chosen for this study would export the expressed protein directly into the culture media. This approach could provide several advantages over the *E. coli* systems traditionally used. Firstly, the CYP17 cDNA sequence would not have to be modified. In *S. cerevisiae* the CYP17 cDNA is fused to the α-factor secretory signal sequence which, during its direction through the yeast's secretory pathway (23, 24), should be translationally cleaved thereby producing a full-length, unmodified P45017α protein. Secondly, there would be no requirement for the optimisation of a preferred translational initiation context at the amino-terminal of the MFα-P45017α fusion protein, as the ATG site of α-factor is in context. Thirdly, the formation of an unwanted secondary structure around the ribosome binding site and initiation site would be reduced, which could lead to increased translation (25). Finally, exporting the expressed P45017α protein directly into the culture media may improve the recovery and yield, since no cell disruption and isolation of P450 enzymes from bacterial membrane fractions would be necessary.

For this study the commercially available yeast amino-terminal FLAG expression system was used. The FLAG octapeptide epitope, linked to the P45017α coding sequence as a fusion protein, it is useful for the rapid affinity purification and facile immunological detection with the anti-FLAG M1 affinity gel and anti-FLAG M2 monoclonal antibodies, respectively. The FLAG tag also contains a five amino acid recognition sequence for enterokinase, thus facilitating the recovery of an intact protein following its proteolytic removal.

We have constructed a novel recombinant FLAG-P45017α fusion protein. The CYP17 cDNA with a nucleotide sequence encoding six histidine residues at the carboxyl-terminus was incorporated into the YEpFLAG-1 expression vector. This construct was inserted downstream from α-factor secretory signal sequence and the FLAG marker

Stellenbosch University http://scholar.sun.ac.za

octapeptide. Sequencing (results not shown), restriction digests, PCR amplification and Southern blot analyses confirmed the correct cloning of the CYP17 into pFLAG-1 to yield pFLAG-CYP17. The cloned expression vector pFLAG-CYP17 was subsequently transformed into a protease deficient *S. cerevisiae* host strain and screened by means of tryptophane selectivity and PCR.

To optimize extracellular excretion from the yeast under the control of the alcohol dehydrogenase promoter (*ADH2*), six different expression media were investigated. Although Northern analyses confirmed transcription of CYP17, immunological investigation of the extracellular, intracellular and periplasmic fractions did not indicate any P45017α protein.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 1 : CLONING AND CONSTRUCTION OF THE FULL-LENGTH HUMAN CYP17 INTO YEpFLAG EXPRESSION VECTOR AND TRANSFORMATION IN *S. CEREVISIAE*

1.1 Materials and Methods

The *E. coli* strain HB101 (Promega) was used in all DNA manipulations. For routine growth in *E. coli*, sterile Luria-Bertani (LB) medium (1% Bacto-Tryptone, 0.5% Bacto Yeast Extract, 1% NaCl, pH 7.0) was used. All media components were purchased from Difco. Selection of bacteria transformants was carried out on LB-agar (15 g/L) containing ampicillin (100 μg/ml). For blue /white selection X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside), 20 mg/ml, was added to LB-agar.

S. cerevisiae strain BJ3505 (a pep4::HIS3 prb-Δ1.6R HIS3 lys2-208 trp1-Δ101 ura3-52 gal2 can1) (Sigma, St. Louis, USA), used for the expression of human cytochrome P45017α, were grown, transformed and analysed according to the manufacturer's instructions. For the routine growth of yeast, YPD media (1% yeast extract, 2% peptone, and 2% glucose) was used. Selection of yeast transformants was carried out in synthetic complete medium containing 0.67% yeast nitrogen base without amino acids, 0.072% amino acids without tryptophane and 2% glucose (SC^{-TRP}). For heterologous expression, *S. cerevisiae* was grown in buffered YPHSM medium (0.1 M potassium phosphate pH 6.0, 1% glucose, 3% glycerol, 1% yeast extract, 2% peptone and 20 mM calcium chloride) or in YPEM medium (0.1 M potassium phosphate pH 6.0, 1% glucose, 3% glycerol, 1% yeast extract, 8% peptone). For the optimisation of expression YEPM media was used, replacing glucose and glycerol with either 2% raffinose and 2% ethanol or 0.8% galactose and 3% glycerol.

Stellenbosch University http://scholar.sun.ac.za

1.1.1 Cloning and construction of the pFLAG-CYP17 expression vector

1.1.1.1 **Polymerase chain reaction (PCR)**

All DNA manipulations were carried out using standard procedures (26). The full-length CYP17, located in the pCD17 vector[1] was used as the template to amplify a full-length CYP17 cDNA. The upstream primer STE1 (5'-ATGCGAATTCATGTGGGAGCTC GTGGCTCTC-3') and the downstream primer 574 (5'-TAGAAGCTT***ATGGTGATG GTGATG***GGTGCTACCCTCAGCCTG-3') were used. The STEI primer introduced an *Eco*RI site (underlined) and the latter primer a nucleotide sequence encoding five histidine residues (bold).

Each PCR amplification mixture (100 μl) contained pCD17 vector DNA (400 ng), 200 μM each of dGTP, dATP, dTTP and dCTP, 0.6 μM each of the sense and antisense primers and 2.5 U proof-reading enzyme Pwo-polymerase (Roche) in Pwo buffer containing $MgSO_4$ (20 mM). The following procedure was carried out in a PCR-Sprint thermo cycler (Hybaid): (1) denaturation of the template at 94°C for 3 min; (2) 35 cycles of denaturing at 94°C for 30 s, annealing at 60°C for 30 s and extension at 72°C for 3 min; (3) and a final extension at 72°C for 15 min. The amplified DNA fragment was purified using an 0.8% low-melting agarose gel in TAE buffer (40 mM) (Tris-HCl pH 8.0, glacial acetic acid and 5 mM EDTA). The DNA samples (1 μg), diluted in 0.2% loading buffer (0.1% Orange G (w/v), 20% Ficoll (w/v), 10 mM EDTA, pH 7.0), were electropheresed (EC Minicell® EC370M), at 55 V for 90 min, stained in ethidium bromide (10 μg/μl) and visualised under a UV lamp (Vilber Lourmat). The CYP17 fragment was excised and purified using the Wizard™ PCR Preps DNA Purification system (Promega, Madison, USA), according to the manufacturer's instructions. The DNA was eluted with deionised water, 50 μl, and stored at –20°C.

1.1.1.2 **Restriction digestion**

The PCR product was digested with *Eco*RI yielding CYP17 cDNA fragments containing *Eco*RI compatible 5'-termini and blunt-end 3'-termini (Figure 1.1) which was

[1] The pCD17 vector was a kind gift from Prof RW Estabrook, UT Southwestern Medical Centre, Dallas Texas, USA.

Stellenbosch University http://scholar.sun.ac.za

subsequently cloned into the YEpFLAG-1 expression vector[2], linearised with *Eco*RI and *Sma*I (Promega, Madison, USA) (Figure 1.2).

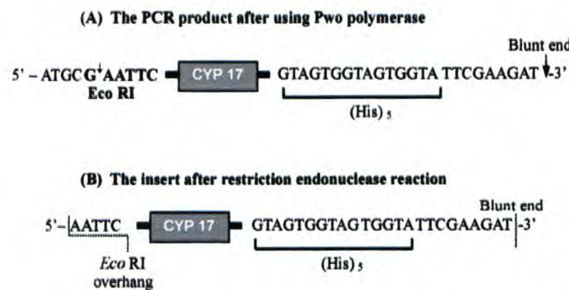

Figure 1.1 PCR amplification (A) and restriction digestion (B) products. The PCR method amplified CYP17 together with the nucleotide sequence encoding five histidine residues and the restriction sites. Pwo proofreading polymerase ensured that a blunt was created (A). The PCR product (A) was digested with *Eco*RI resulting in a protruding *Eco*RI 5' end and a blunt 3' end (B).

Two separate restriction digestions, one containing amplified CYP17 cDNA and the other YEpFLAG-1 DNA were carried out. The amplified CYP17 cDNA (2 μg) was digested at 37°C for 90 min, in a final volume of 30 μl containing 5 μl restriction buffer B (supplied), acetylated BSA (5 μg) and *Eco*RI (10 U)(Promega). The YEpFLAG expression vector was digested with *Eco*RI and *Sma*I (Promega) in two consecutive steps. The plasmid DNA (1 μg) was digested at 25°C for 60 min, in a final volume of 20 μl containing 2 μl Multicore buffer (supplied), acetylated BSA (2 μg) and *Sma*I (5 U)(Promega). The plasmid DNA was further digested at 37°C for 60 min, in a final volume of 30 μl containing 1 μl Multicore buffer, acetylated BSA (2 μg) and *Eco*RI (5 U). The digested CYP17 fragment and linearised vector were electrophoresed on an 0.8% low-melting agarose gel in TAE buffer (40 mM) and purified as described in 1.1.1.1.

1.1.1.3 Ligation

The purified CYP17 cDNA was subsequently ligated using 1:1 and 1:3 molar ratio of vector:insert DNA. For a 1:1 molar ratio, DNA insert (11 ng) and vector (50 ng) was

[2] Kindly donated by Prof. W. H. van Zyl of the Microbiology Department, University of Stellenbosch, Stellenbosch, South Africa.

Stellenbosch University http://scholar.sun.ac.za

used in a 10 µl ligation reaction containing T4 ligase (1.5 U)(Promega), 0.5 µl ligase buffer (30 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 10 mM DTT and 1 mM ATP) and ATP (0.5 mM). The reaction was incubated at 16°C for 18 hours and the construct stored at –20°C.

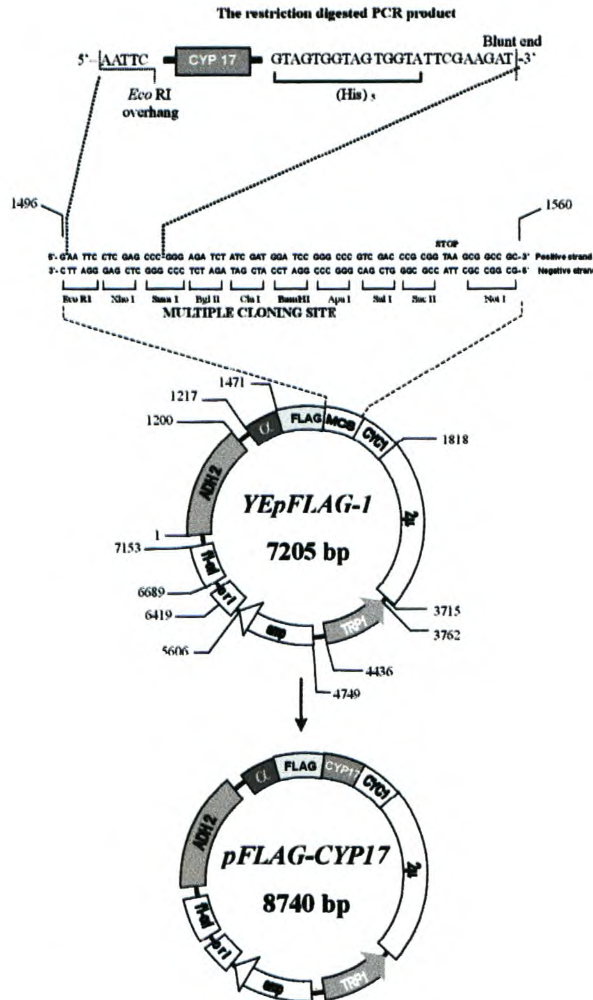

Figure 1.2 Schematic representation of the cloning strategy for the construction of the pFLAG-CYP17 expression vector. The PCR amplified CYP17 was digested with *Eco*RI and inserted into the YEpFLAG-1 expression vector. The CYP17 was cloned in reading frame of the α-factor secretion signal and contained nucleotide bases encoding five histidine residues at the 5'-terminal.

1.1.1.4 Transformation

Competent bacteria (HB101) were prepared and the transformation was carried out as described in protocol 1 (26). Competent bacterial cell suspension (50 µl) was added to 5 µl of the ligation reactions, mixed gently and placed in ice for 30 min. The mixtures

Stellenbosch University http://scholar.sun.ac.za

were subsequently heat shocked at 42°C for 90 s and incubated in ice for 90 s. After the addition of 800 μl SOC medium (37°C) (2% tryptone, 0.5% yeast extract and 0.05% NaCl, 250 mM KCl, pH 7.0, 10 mM $MgCl_2$ and 20 mM glucose), the cultures were incubated at 37°C for 45 min in a shaking incubator at 225 rpm. The cell suspension (100 and 150 μl) was plated out on LB-agar medium containing ampicillin (100 μg/ml) and incubated at 37°C overnight.

1.1.2 Analyses of *E. coli* transformants

Positive colonies were subsequently incubated overnight in LB medium (5 ml) at 37°C and the plasmid DNA was isolated using the Wizard™ *Plus* Minipreps DNA Purification System (Promega), according to the manufacturer's instructions. The purified recombinant plasmid DNA was digested with *Eco*RI and *Bam*HI and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit (Macherey-Nagel) as described in the manual for high-copy number plasmids. The purified isolated pFLAG-CYP17 plasmid DNA was further analysed by PCR, DNA sequencing and Southern blotting.

1.1.2.1 **PCR amplification of *E. coli* transformants**

The pFLAG-CYP17 vector was analysed by PCR using the forward primer YαN-21 (5'-AGCACAAATAACGGGTTATTG-3') and reverse primer YcC-21 (5-TACAGACGC GTGTACGCATGT-3'), supplied with the FLAG Expression System kit (Sigma, St. Louis, USA). Each PCR amplification mixture (50 μl) contained pFLAG-CYP17 DNA (100 ng), 200 μM each of dATP, dTTP, dCTP and dGTP, 2.5 μM each of the YαN-21 and YcC-21 primers and DNA Taq polymerase (2.5 U) (Promega, Madison, USA) in Tris-HCl (10 mM), pH 9.0, containing 50 mM KCl, 0.1% v/v Triton X-100 and 1.5 mM $MgCl_2$. The following procedure was carried out in a MiniCycler™ (MJ Research): (1) denaturation at 94°C for 3 min; (2) 35 cycles of denaturing at 94°C for 30 s, annealing at 60°C for 30 s and extension at 72°C for 3 min; (3) and a final extension at 72°C for 15 min. The PCR product was analysed by agarose gel electrophoresis.

Stellenbosch University http://scholar.sun.ac.za

1.1.2.2 DNA sequencing strategy of *E. coli* transformants

The purified pFLAG-CYP17 plasmid (100 ng/µl), together with its insertion sites was analysed by DNA sequencing using eight primers (1.1 pmol/µl) (Figure 1.3). The sense primers were STE1; 4011 (5'-TGGCAGCTCATCGAAGGCT-3'); 1496 (5'-GAAGC TCTACGAGGAGATTGACCAG-3') and the antisense primers were 3989 (5'-AGCCTTCGATGCAGCTGCCAG-3'); 8982 (5'-CTGGTCAATCTCCTCGTAGAGCT TC-3'); STE2 (5'-AGCTAAGCTTTTAGGTGCTACCCTCAGCCTG-3'). Upstream and downstream sequencing primers (YαN-21 and YcC-21) were also used for the sequencing reaction. Nucleotide sequences were determined using the Bigdye™ Version 2 determinator sequencing kit (model 373A ABI, Applied Biosystems, Foster City, CA).

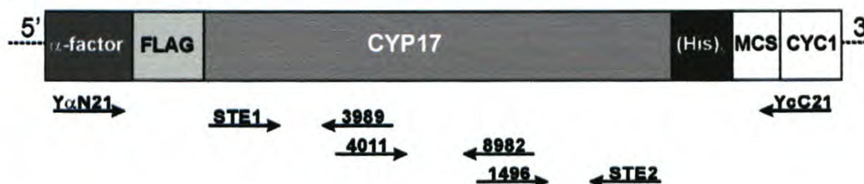

Figure 1.3 Schematic representation of the DNA sequencing strategy for the full-length human CYP17 in the pFLAG-CYP17 expression vector. Primers YαN21 and YcC21 sequence from the outside of the flanking regions of the CYP17 insert confirmed correct insertion.

1.1.2.3 Southern hybridisation of CYP17 DNA

The recombinant pFLAG-CYP17 plasmid DNA, previously linearised *Eco*RI, was transferred onto a nylon Hybond-N^+membrane (Amersham, High Wycombe, UK) as described by Sambrook *et al.*, Section 10.46 (26). The plasmid DNA was fixed and the membrane hybridised with DIG-labelled (digoxigenin-labelled) human CYP17 cDNA[3] (25 ng/ml), as described in the DIG High Prime DNA Labelling and Detection kit (Roche, Mannheim, Germany). Immunodetection of the hybrids were carried out with anti-digoxigenin-AP Fab fragments (1:5000) and subsequently visualised with colorimetric substrates NBT/BCIP, according to the manufacturer's instructions (Roche, Mannheim, Germany).

---

[3] The DIG-labelled human CYP17 cDNA was prepared and kindly donated by Dr. A. C. Swart, Department of Biochemistry, University of Stellenbosch, Stellenbosch, South Africa.

Stellenbosch University http://scholar.sun.ac.za

1.1.3 *S. cerevisiae* transformation and screening of transformants

S. cerevisiae strain BJ3505 was transformed with the pFLAG-CYP17 expression vector using the LiOAc/DMSO method (27). Three controls, viz. the original YEpFLAG-1 plasmid without insert, a negative control without plasmid DNA and a positive control, YEpFLAG-1 BAP were included. Transformed cells were grown on SC^{-TRP}. The yeast DNA was isolated from positive colonies according to the rapid yeast DNA preparation protocol described by Hoffman and Winston (28) and CYP17 was amplified using the upstream primer STE1 and the downstream primer 574. Each PCR amplification mixture (50 µl) contained pFLAG-CYP17 DNA (100 ng) and the reaction was carried out as described in section 1.1.2.1. The amplified DNA was analysed by agarose gel electrophoresis.

Single colonies containing verified recombinants in BJ3505 were grown in 10 ml SC^{-TRP} overnight at 30°C. The cells were isolated by centrifugation at 3000 × g for 5 min at room temperature and resuspended in fresh SC^{-TRP} containing 15% glycerol. Cells were frozen in liquid nitrogen and stored at –80°C.

1.2 Results

1.2.1 Confirmation of the integrity of the YEpFLAG-1 plasmid

The authenticity of the donated YEpFLAG-1 plasmid, was confirmed by restriction digest analyses with *Eco*RI/*Sma*I followed by 1% agarose gel electrophoresis. The linearised plasmid DNA yielded the expected band corresponding to 7205 bp as indicated in Figure 1.4, lane 3.

1.2.2 Analyses of *E. coli* transformants

Transformed bacterial colonies were grown in small-scale cultures. The plasmid DNA isolated from each culture was digested with *Eco*RI and *Bam*HI and analysed by 1% agarose gel electrophoresis. The pFLAG-CYP17 plasmid (8740 bp), containing the insert, yielded two 7175 and 1565 bp fragments as shown in Figure 1.5.

Stellenbosch University http://scholar.sun.ac.za

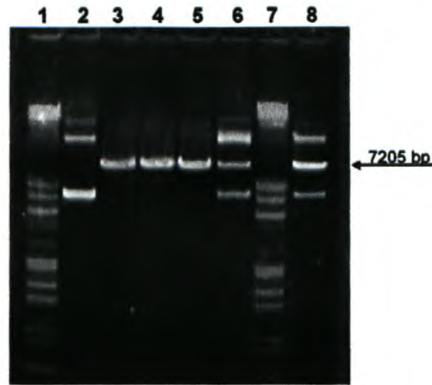

Figure 1.4 Restriction digest analysis of the original YEpFLAG-1 plasmid. Lanes 1 and 7: DNA marker, λDNA/*Eco*RI/*Hind*III. **Lane 2:** undigested YEpFLAG-1 DNA. **Lane 3:** YEpFLAG-1 DNA (0.5 μg) digested with *Eco*RI and *Sma*I. **Lane 4:** *Eco*RI digest of YEpFLAG-1 DNA (0.5 μg). **Lane 5:** YEpFLAG-1 DNA (0.5 μg) with *Sma*I. **Lane 6:** YEpFLAG-1 DNA (0.5 μg) digested with *Bam*HI. **Lane 8:** YEpFLAG-1 DNA digested with *Sma*I.

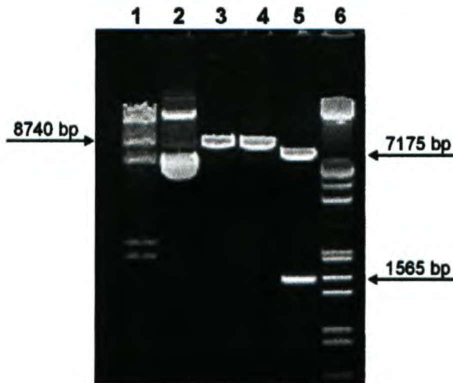

Figure 1.5 Restriction digest analysis of pFLAG-1 plasmids containing the CYP17. Lane 1: DNA marker, λDNA/*Hind*III. **Lane 2:** Undigested pFLAG-CYP17 plasmid (1 μg). **Lane 3:** *Eco*RI digest of pFLAG-CYP17 DNA (0.5 μg) containing the CYP17 yielding an 8740 bp fragment. **Lane 4:** *Bam*HI digest of pFLAG-CYP17 DNA (0.5 μg). **Lane 5:** *Eco*RI/*Bam*HI digest of plasmid DNA (0.5 μg) containing the CYP17 yielding 7175 and 1565 bp fragments. **Lane 6:** DNA marker, λDNA/*Eco*RI/*Hind*III.

Plasmid DNA isolated from each culture was amplified with CYP17 specific primers and analysed by 1% agarose gel electrophoresis as shown in Figure 1.6. PCR and DNA sequencing analyses confirmed that the CYP17 fragment had been correctly inserted in reading frame in the YEpFLAG-1 vector. A 1793 bp fragment was amplified using primers (YαN-21 and YcC-21) which recognise the sequence flanking the inserted DNA shown in Figure 1.6. These primers were subsequently used to sequence 300 bp from the 5' and 3' end of the CYP17 cDNA. The integrity of CYP17 (1193 bp) was confirmed by further sequencing and showed that no mutations or alterations occurred during the cloning procedures.

Stellenbosch University http://scholar.sun.ac.za

Figure 1.6 PCR amplification analysis of the recombinant YEpFLAG-CYP17 plasmids. Lane 1: Negative control, PCR amplification without DNA template. **Lane 2:** DNA marker, λDNA /*Hind*III. **Lanes 3 and 4:** Amplified CYP17 fragment, 1793 bp, using primers YαN-21 and YαC-21. Each lane contained 0.5 μg DNA. **Lane 5:** DNA marker, λDNA/*Eco*RI/*Hind*III.

Southern blot analysis was carried out using a DIG-labelled probe, DIG-CYP17, complementary to the human CYP17 cDNA sequence. The linearised pFLAG-CYP17 recombinant plasmid was positively identified as shown in Figure 1.7.

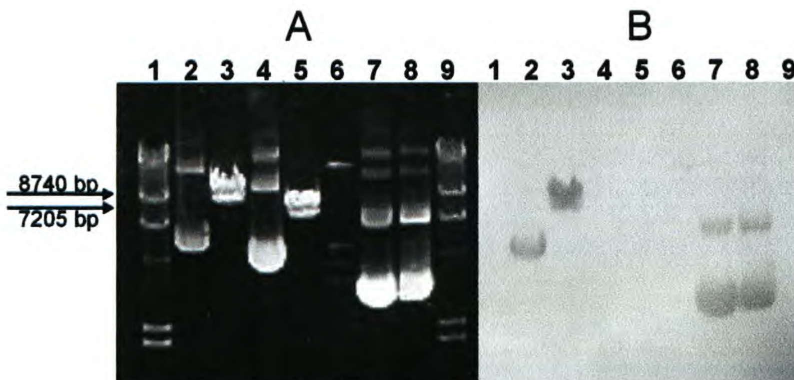

Figure 1.7 Southern blot analysis. Lanes 1 and 9: DNA marker , λDNA /*Hind*III. **Lanes 2 and 3:** Undigested and *Eco*RI digest of the recombinant pFLAG-CYP17 plasmid, respectively. **Lanes 4 and 5:** Undigested and *Eco*RI digest of the original YEpFLAG-1 plasmid, respectively. **Lane 6:** DNA marker, λDNA /*Hind*III/ *Eco*RI II. **Lanes 7 and 8:** Undigested and *Eco*RI digested pCD17α plasmid, respectively

The probe hybridised to the original plasmid (pCD-CYP17) from which the CYP17 cDNA was amplified. No hybridisation was detected with the plasmid YEpFLAG-1 containing no CYP17 insert.

1.2.3 Analyses of *S. cerevisiae* transformants

Small-scale cultures of the transformed yeast cells were grown and total yeast DNA was isolated. DNA purified from two cultures was amplified with gene specific primers

Stellenbosch University http://scholar.sun.ac.za

(STE1 and 574) and the PCR products analysed by 1% agarose gel electrophoresis. The cloned pFLAG-CYP17 vector should yield one amplified CYP17 fragment (1549 bp) as shown in Figure 1.8.

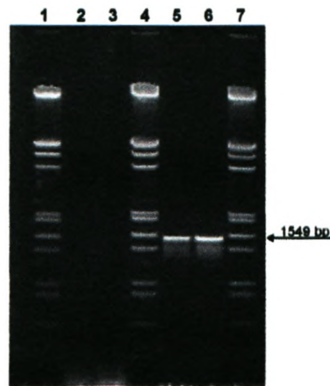

Figure 1.8 PCR amplification analysis of pFLAG-CYP17 plasmid DNA isolated from BJ3505 yeast cells. Lanes 1, 4 and 7: DNA marker, λDNA /*Eco*RI/*Hind*III. **Lanes 2 and 3:** Negative controls, PCR amplifications without DNA template. **Lanes 5 and 6:** Amplified CYP17 fragment, 1549 bp. Each lane contained 0.5 μg DNA.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 2 : CULTURE OF S. *CEREVISIAE* EXPRESSING P45017α AND OPTIMISATION OF PROTEIN EXPRESSION

2.1 Materials and Methods

2.1.1 Growth and optimisation of *S. cerevisiae* P45017α expression

The heterologous expression was carried out according to the FLAG Expression System protocol (Sigma, St. Louis, USA) and is summarised in Table 2.1. Growth media (SC^{-TRP}), 25 ml, were inoculated with colonies containing pFLAG-CYP17 and control plasmids (strain BJ3505) and incubated at 30°C in a shaking incubator, 175 rpm, for 48 hours. At selected time intervals (24 and 48 hours) 1 ml samples were collected and the cell density determined spectrophotometrically at 600 nm. After 48 hours different expression media, 300 ml, (Table 2.1) were inoculated with 20 ml of cell suspension and incubated at 30°C in a shaking incubator, 175 rpm, for 162 hours. Attempting to optimise expression, cells from a single culture, flask 5, were harvested after 48 hours by

Table 2.1 Expression media used in the optimisation of CYP17 expression at transcriptional level. The expression method, according to the manufacturer's protocol, was carried out in flask 3 and 4 while optimisation was carried out in flask 5 to 8.

	Flasks with different expression media							
	YEpFLAG-1	WT[1]	pFLAG-CYP17					
Flask:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Growth media:	SC^{-TRP} [2] 2% glucose	SC^{-TRP} 2% glucose	SC^{-TRP} 2% glucose	SC^{-TRP} 2% glucose	SC^{-TRP} 8% glucose	SC^{-TRP} 2% glucose	SC^{-TRP} 8% glucose	SC^{-TRP} 8% glucose
Expression media:	YPHSM[3] + 1% glucose 3% glycerol	YPHSM + 1% glucose 3% glycerol	YPHSM + 1% glucose 3% glycerol	YPEM[4] + 1% glucose 3% glycerol	YPEM + 2% raffinose 2% ethanol	YPEM + 2% raffinose 2% ethanol	YPEM + 0.8% galactose 3% glycerol	SC^{-TRP} + 0.8% galactose

---

[1] WT: wild type strain BJ3505

[2] SC^{-TRP}: synthetic complete medium minus tryptophane

[3] YPHSM: Yeast Peptone High Stability Media (0.1 M potassium phosphate pH 6, 01 % yeast extract, 8 % peptone and 20 mM calcium chloride)

[4] YPEM: Yeast Peptone Expression Media (0.1 M potassium phosphate pH 6, 01 % yeast extract, 2 % peptone)

Stellenbosch University http://scholar.sun.ac.za

centrifugation at 2000 × g for 5 min at room temperature. The cell pellet was resuspended in 10 ml of ddH_2O and centrifuged at 2000 × g for 5 min at room temperature. The washed cells were resuspended in expression medium and added to the fresh expression medium.

The growth rate was determined at 4, 18, 24, 72 and 164 hr time intervals by monitoring yeast cell density spectrophotometrically at 600 nm in 1 ml samples. At the stipulated time intervals, 50 ml of the expression culture was harvested by centrifugation at 5000 × g for 5 min at 4°C. The supernatant (medium) was stored at –20°C and the cells were immediately used to determine RNA expression levels and intracellular and periplasmic protein expression. The supernatant was used to analyse extracellular protein expression levels

2.1.2 Evaluation of P45017α expression at transcriptional level

2.1.2.1 **RNA isolation**

Total RNA was isolated using the phenol/chloroform/isoamyl alcohol isolation reagent according to the methods described by Hoffman *et al.* in Current Protocols, section 13.2.2 (29).

The harvested yeast cells were thawed on ice and resuspended in 300 μl cold STE buffer (100 mM NaCl, 10 mM Tris-HCl pH 8.0 and 1 mM EDTA pH 8.0). The resuspended cells were transferred to ice-cold 2 ml microcentrifuge tubes and 300 μl of phenol, 1 volume of acid washed glass beads and 3 μl of SDS (20% w/v) was added. The suspension was vortexed for 50 s, 300 μl of a chloroform: isoamyl alcohol (24:1 v/v) mixture was added and vortexed for a further 20 s. The broken yeast cells were centrifuged (Hettich Microliter Type 2021) in the same 2 ml microcentrifuge tubes at 13000 × g for 10 min at 4°C. The upper clear layers (200-250 μl) were carefully pipetted into sterile, ice-cold 2 ml microcentrifuge tubes. The tubes were briefly vortexed after adding 1 volume of PCI solution (phenol: chloroform: isoamyl alcohol at 25: 24: 1 (v/v/v)) to each tube, and centrifuged at 13000 × g for 5 min at 4°C. The upper clear layers were removed and transferred into new 2 ml microcentrifuge tubes. The

Stellenbosch University http://scholar.sun.ac.za

extraction procedure was repeated once more with one volume of PCI solution and the upper clear layers removed and transferred to new ice-cold 2 ml microcentrifuge tubes. RNA was precipitated with 0.1 volume NaCl (5M NaCl, vortexed twice) and two volumes of 100% ethanol. The tubes were placed at –80°C for 4 hours and the RNA was subsequently collected by centrifugation at 13000 × g for 10 min at 4°C. The RNA pellets were dried slightly at 65°C for 10 min and resuspended in 50 μl of distilled water. The concentration of the RNA was determined spectrophotometrically at 260 nm RNA samples were stored at –80°C and dot blot and Northern blot analyses were subsequently carried out.

2.1.2.2 Northern hybridisation of CYP17 mRNA

The RNA preparations (10 μg) were denatured, electrophoresed on a 0.8% agarose gel, containing 37% (v/v) formaldehyde, and transferred onto a Hybond$^+$ nylon membrane (Amersham) by capillary transfer as described by Sambrook *et al.*, section 10.46 (26). The nylon membrane was pre-hybridised in 30 ml high SDS hybridisation buffer (7% SDS, 50% formamide, 5 × SSC, 2% blocking reagent, 50 mM sodium-phosphate, pH 7.0 and 0.1% N-lauroylsarcosine) at 50°C for 6 hours. The pre-hybridisation buffer was discarded and the DIG-labelled CYP17 DNA probe (120 ng/μl), diluted in the hybridisation buffer and heat-denatured by boiling in a waterbath for 10 min, was added. After hybridising overnight in a shaking hybridisation bath, the membrane was washed twice in 2 × wash solution (2 × SSC, 0.1% SDS) for 15 min at room temperature and twice in 0.5 × wash solution (0.5 × SSC, 0.1% SDS) for 15 min at 50°C.

The hybrids were immunodetected with anti-digoxigenin-AP Fab fragments (1:5000) and then visualised with the colorimetric substrates NBT/BCIP, according to the manufacturer's instructions (Roche, Mannheim, Germany).

2.1.2.3 RNA dot blot

Equal concentrations of isolated total RNA obtained from culture samples at specific time points, with were spotted onto a dry membrane. The RNA spots were dried for 15 min. at room temperature and cross-linked to the membrane by UV for 6 s. Northern blot analysis was carried out as described in section 2.1.2.2.

Stellenbosch University http://scholar.sun.ac.za

2.1.3 P45017α protein expression

2.1.3.1 **Extracellular protein expression**

Yeast cultures (15 ml) obtained at selected time intervals (4, 24, 48, 72, 96 hours) were harvested by centrifugation at 2000 × g for 5 min at room temperature. Extracellular protein fractions were prepared by adding PMSF to the supernatant to a final concentration of 1 mM. The fractions were transferred into pre-washed Ultrafree-15 (Biomax™ High Flow Ultrafiltration membrane, Millipore) ultrafiltration units (MCO of 10000) and centrifuged at 2000 × g for 2 hours at 4°C in a fixed angle rotor. The fractions (30 × concentrated) were transferred into 1.5 ml eppendorf tubes and stored at 20°C. The expressed proteins in the concentrated fractions were subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot analyses.

2.1.3.2 **Preparation of the intracellular protein fraction**

Yeast cells were harvested at 72 hours by centrifugation at 5000 × g for 5 min. at 4°C. The resulting pellet was resuspended in 5 ml of protein extraction buffer (200 mM Tris-HCl pH 8.0, 400 mM ammonium sulphate, 10 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 1 mM PMSF and 7 mM β-mercaptoethanol). The resuspended cells were transferred to 50 ml centrifuge tubes and centrifuged at 5500 × g for 10 min. at 4°C. The cells were resuspended in protein extraction buffer (400 μl), transferred to ice-cold 2 ml microcentrifuge tubes containing acid-washed glass beads (300 μl). The cells were mechanically disrupted by vortexing 20 times for 30 s with 30 s intervals on ice. The breakage of the yeast cells was monitored under a microscope. After incubation on ice for 30 min, the broken cells were centrifuged at 13500 × g for 5 min at 4°C. The supernatant was transferred to 2 ml microcentrifuge tubes and centrifuged at 13500 × g for 60 min at 4°C. The proteins in the supernatants were precipitated by the addition of two volumes of acetone (100% v/v) followed by thorough mixing and incubation for 60 min at –20°C. The contents of the tubes was gently mixed at specific intervals. The precipitated proteins were collected by centrifugation at 13500 × g for 10 min at 4°C and the supernatant discarded. The protein pellet was resuspended in ice-cold HEPES buffer (240 μl)(10 × HEPES buffer, 7 mM β-mercaptoethanol, 1.25 mM PMSF) and glycerol

Stellenbosch University http://scholar.sun.ac.za

(80% v/v) (75 μl). Isolated intracellular proteins were stored at −80°C until and subsequently subjected to Western blot analysis.

2.1.3.3 **Preparation of the periplasmic and intracellular protein fractions**

P45017α retained in the periplasmic space were isolated by spheroplasting as described by La Grange *et al.* (30). Yeast cells harvested at 72 hours by centrifugation at 5000 × g for 5 min. at 4°C were washed in ice-cold distilled water and resuspended in 1.5 ml ice-cold Tris-HCl (50 mM, pH 7.4). The resuspended cells (300 μl) were added to 1.5 ml phosphate buffer (6.7 mM K_2HPO_4 pH 7.5, 120 mM β-mercaptoethanol, 1.25 mM PMSF) and distilled water (1.2 ml) incubated for 30 min at room temperature. The cells were lysed with Quantazyme (Quantum Biotechnologies) (20 U/μl), incubated for 90 min at room temperature and centrifuged at 1000 × g for 1 min at 4 °C.

The supernatants, representing the proteins in the periplasmic fraction, were precipitated overnight with two volumes of ice-cold acetone (100% v/v), harvested by centrifugation and resuspended in 50 μl HEPES buffer (20 mM HEPES pH 8.0, 5 mM EDTA pH 8.0, 7 mM β-mercaptoethanol, 1.25 mM PMSF). The periplasmic fractions were subjected to SDS-PAGE and Western blot analyses.

The pellets, representing spheroplasts, were washed twice with ice-cold phosphate buffer and once with sorbitol solution (1 M sorbitol, 0.2% v/v $CaCl_2$). The spheroplasts were centrifuged at 1000 × g for 1 min at 4°C and the pellet was subsequently resuspended in distilled water (100 μl), lysed by adding 10 μl SDS (20% w/v) and incubated for 10 min at 65°C. The intracellular proteins were harvested by centrifugation at 1000 × g for 1 min at 4°C and precipitated with two volumes of ice-cold acetone and stored at −20°C. The intracellular fractions were subsequently analysed by SDS-PAGE and Western blot.

2.1.4 SDS-PAGE and Western blot analyses

SDS-polyacrylamide gel electrophoresis was carried out according to the method of Laemmli (31) using an acrylamide concentration of 10% (v/v). Fractions to be analysed were denatured in equal volumes of treatment buffer (0.125 M Tris-HCl pH 6.8, 4%

Stellenbosch University http://scholar.sun.ac.za

SDS, 20% glycerol and 10% β-mercaptoethanol) and 20 μl bromophenol blue, boiled for 5 min and immediately placed on ice. Low Range molecular weight markers (Biorad) and Rainbow coloured protein molecular weight markers (Amersham Pharmacia) consisting of myosin (220 kDa), phosphorylase b (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (21 kDa) and lysozyme (14.4 kDa) were used.

Electroblotting was carried out in a wet blotting system. Electrophoretically resolved protein samples were transferred onto a nitrocellulose membrane (Schleicher&Schueller) and the primary antigen-probe, a monoclonal anti-FLAG M_2 mouse antibody (Sigma, St. Louis, USA) was visualised by peroxidase staining colorimetric method. The probe was diluted at 1:1000 in casein buffer. A FLAG-BAP (bacterial alkaline phosphatase), supplied with the kit, was protein used as a positive control.

2.2 Results

2.2.1 Growth of *S. cerevisiae*

Cell culture samples were taken during the expression phase at 4, 18, 24, 72 and 164 hours time intervals to monitor the cell growth rate. Results shown in Figure 2.1 indicate that cells grew linearly up to 72 hours where after growth declined due to autolysis of the yeast cells.

2.2.2 Transcription levels of CYP17 in yeast during induction phase

The same samples were analysed to ascertain the level of CYP17 transcription and results showed that the highest level of total RNA synthesis was between 18 to 24 hours (Figure 2.2). Total RNA isolated from BJ3505 cells transformed with pFLAG-CYP17 was subsequently analysed by RNA dot blot analysis using the DIG-labelled CYP17 cDNA as the probe. As shown in Figure 2.3, RNA could only be detected at 18-24 hours. Cells grown in raffinose (2% w/v) as the carbon source and ethanol (2% v/v) as non-fermentable carbon source (flask 5 and 6), produced the highest levels of CYP17 transcripts. Cells grown in the original

Stellenbosch University http://scholar.sun.ac.za

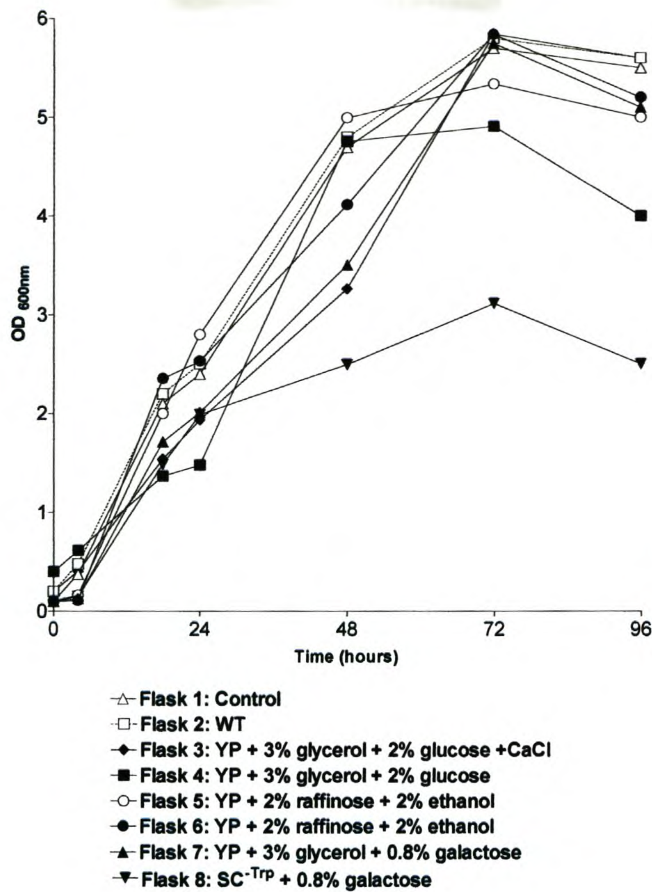

Figure 2.1 Growth curves of BJ3505 transformants in the expression phase. YP: yeast peptone, SC^{-Trp}: selective minimal medium without tryptophane.

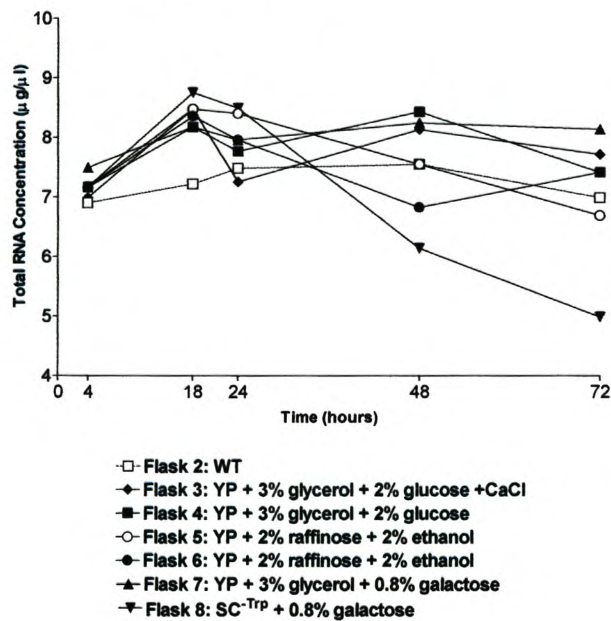

Figure 2.2: Total RNA in yeast cultures assayed at 4 – 72 hours. YP: yeast peptone, WT: wild-type, SC^{-Trp}: synthetic complete medium without tryptophane.

Stellenbosch University http://scholar.sun.ac.za

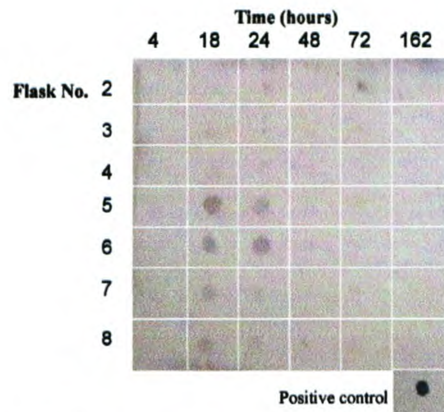

Figure 2.3 Dot blot analysis of the CYP17 transcripts expressed in BJ3505 assayed in different expression media. Total cellular RNA, 35µg, from flasks 2-8, collected at 4, 18, 24, 48, 72 and 162 hours, was spotted in each lane. DIG-labelled CYP17 DNA was used as a probe. The intensity of the coloured blot reflects the concentration of the CYP17 mRNA in the preparations.

expression media with glycerol as the non-fermentable carbon source, did not show significant amounts of RNA (flasks 3 and 4).

Samples, taken at 24 hours, subjected to Northern blot analyses, previously shown to have the highest mRNA, confirmed full-length CYP17 transcripts without premature termination of the transcription (Figure 2.4). The relative position of 18S and 28S ribosomal RNA fractions indicate the estimated size of CYP17 synthesised in yeast to be approximately 2.2 kb. This would be consistent with the predicted DNA fragment size

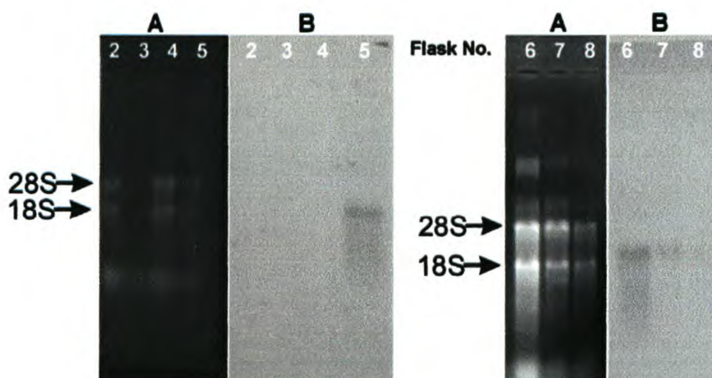

Figure 2.4: Analyses of CYP17 mRNA from yeast strain BJ3505 transformed with pFLAG-CYP17 grown for 24 hours in specified media. A: Formaldehyde agarose gel and B: Northern blot analysis Isolates of total RNA, 10 µg, were analysed on a agarose gel (1.2%) in formaldehyde/MOPS buffer, in the presence of 0.5 µg/ml ethidium bromide. The number in each lane represents the flask number. DIG-labelled CYP17 cDNA was used as a probe. Yeast ribosomal RNAs are indicated at 18S and 28S.

Stellenbosch University http://scholar.sun.ac.za

obtained if transcription of the CYP17 cDNA in the plasmid construct is initiated and terminated at the *ADH2* promoter and CYC1 terminator sites, respectively.

2.2.3 Intra- and extracellular expression of P45017α

SDS-PAGE analyses of the expressed proteins in the YPHSM medium are shown in Figure 2.5. Yeast transformed with the parent YEpFLAG-1 vector expressed a small amount of background proteins (lanes 2, 3 and 4). The transformed clones, however, produced proteins after 48 hours which appear to have molecular masses ranging from 45 to 60 kDa. The expected molecular mass of the P45017α fusion protein was in the order of 58 kDa.

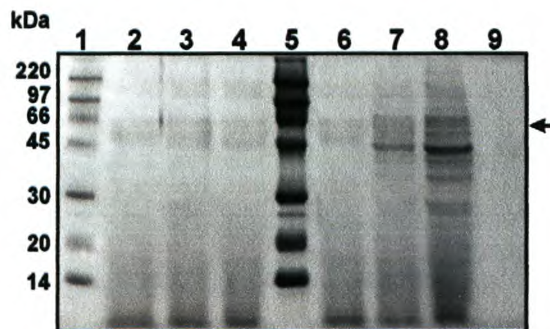

Figure 2.5: SDS-PAGE analysis of the extracellular protein expression in YPSHM medium. Lanes 1 and 5: Low-molecular weight Marker. **Lanes 2, 3 and 4:** Negative control, BJ3505/YEpFLAG-1 cell culture, total protein (10 μg) assayed at 24, 48 and 72 h. **Lanes 6, 7 and 8:** BJ3505/pFLAG-CYP17 cell culture, total protein (10 μg) assayed at 24, 48, and 72 hours. The apparent molecular mass of the FLAG-P45017α fusion protein is 59 kDa indicated by the arrow.

In an attempt to obtain better resolution preparative SDS-PAGE was performed. The larger acrylamide gel (14 ×16 cm) yielded greater resolution in the 46 to 66 kDa range (Figure 2.6).

FLAG-P45017α fusion protein was not detected with Western blot analysis of intracellular (Figure 2.7) and periplasmic (Figure 2.8) proteins expressed by cells transformed with pFLAG-CYP17.

Stellenbosch University http://scholar.sun.ac.za

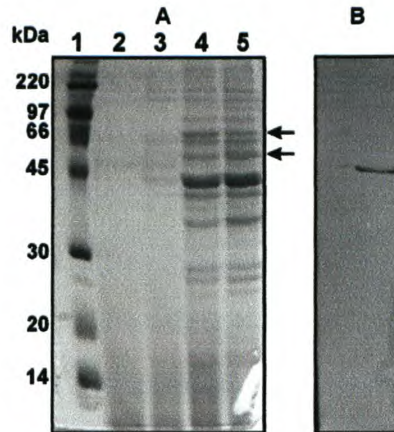

Figure 2.6: Analyses of extracellular protein expression. A: SDS-PAGE analysis: Lane 1: Low-molecular weight Marker. **Lanes 2 and 3:** Negative control, BJ3505/YEpFLAG-1 cell culture, total protein assayed at 24 and 72 h. **Lanes 4 and 5:** BJ3505/pFLAG-CYP17 cell culture, total protein (100 μg) assayed at 24 and 72 hours. Two additional proteins at 58 kDa and 66 kDa detected in the induction phase, are indicated by the arrows. **B Western blot analysis:** Positive control containing BAP protein expressed by BJ3505 transformed with pFLAG-BAP. BAP protein detected with an anti-FLAG M2 monoclonal antibody at 1:1000.

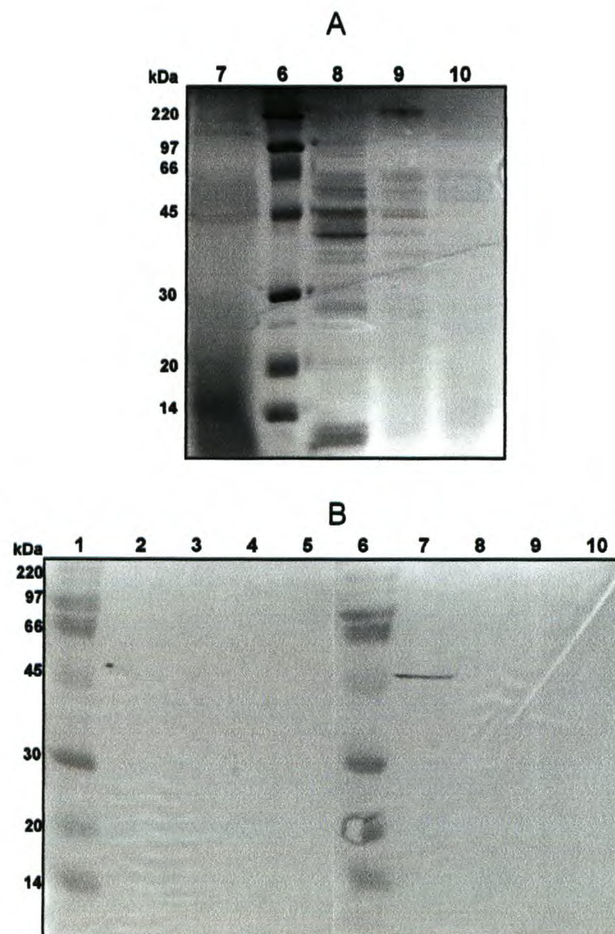

Figure 2.7: Analyses of total intracellular protein expression by transformed yeast cells in specified growth media, assayed at 72 hours A: SDS-PAGE analysis and B: Western blot analysis. Lanes 1 and 5: Low-molecular weight Marker. **Lanes 2 and 7:** Positive control containing BAP protein expressed

Stellenbosch University http://scholar.sun.ac.za

by BJ3505 transformed with pFLAG-BAP. **Lanes 3, 4, 5, 8, 9 and 10:** Total protein (100 μg) in BJ3505/ pFLAG-CYP17 cell cultures grown in flask 2 (lane 3), flask 3 (lane 4), flask 4 (lane 5), flask 5 (lane 8), flask 6 (lane 9) and flask 7 (lane 10). Detection was carried out using an anti-FLAG M2 monoclonal antibody at 1:1000.

Figure 2.8 SDS-PAGE analyses of the intracellular and periplasmic expressed proteins in BJ3505 cells transformed with pFLAG-CYP17, assayed at 72 hours. Lane 1: Low-molecular weight Marker. **Lane 2:** Positive control containing BAP protein **Lanes 3 and 5:** proteins isolated from periplasmic space. **Lane 4:** intracellular proteins (20 μg) isolated from the spheroplasts. No bands were detected in the immunoblot with anti-FLAG antibody in lanes 3, 5 and 6 (results not shown).

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 3 : DISCUSSION

In this Chapter, the cloning of an unmodified human cytochrome CYP17 cDNA into a *S. cerevisiae* expression vector was described. The expression is under the control of the α-factor secretory signal sequence (MFα-factor), which would facilitate the export of P45017α protein into the culture medium. RNA blot experiments revealed that the level of CYP17 mRNA appeared to be higher in yeast cells grown on raffinose and ethanol than on the conventional glycerol source. SDS-PAGE experiments showed that two unidentified proteins were extracellulary expressed and present in the medium, but could however not be characterised by inmmunoblotting experiments.

Since the first report of the successful expression a recombinant cytochrome P450 enzyme in *S. cerevisiae* (32), a steady increase of publications on various heterologously expressed cytochromes P450, originating from a wide range of species from mammals to plants, have been noted as summarised in Part I Chapter 2 Section 2. The use of yeast-based systems has become popular in studying the enzymatic properties of cytochromes P450, as the yeasts contain an endoplasmic reticulum, endogenous NADPH-P450 reductase and cytochrome b5 ensuring a suitable environment for the functional expression of cytochromes P450 *in vivo* (12) or in isolated yeast microsomes (33, 34). Furthermore, yeast strains have been used to elucidate cytochrome P450 structure-function relationships by expressing different forms of P450 enzymes altered by site-directed mutagenesis (35, 36, 37). Although researchers are striving to optimise the specific activity of cytochromes P450 expressed in yeast, either by co-expression (38, 39) or alteration of the amino terminal (40), not many reports concentrate on the yield of cytochromes P450. Early reports have shown that the yeast expression system expresses relatively low levels of 8 nmol/L culture (41), but with newly developed yeast expression systems such as Pichia pastoris (42, 43) and *Schizzosaccharomyces pombe* (44) improvement in expression levels are possible (45, 46).

The aim of this project was the large-scale production of a pure full-length P45017α protein. Since the ultimate goal was to raise polyclonal and monoclonal antibodies against human P45017α, a pure recombinant protein rather than a catalytically active enzyme was required. Therefore, a different approach, the secretion of heterologously

expressed P45017α into the culture medium, was attempted. Directing the expressed protein through the secretory pathway of the yeast ensures correct folding of the protein and protection from cytoplasmic proteases, which could ultimately result in greater stability and purity of the P45017α, simplifying the purification procedure.

Transformed yeast cells were initially grown on YPHSM and YPEM media. The former was the preferred expression medium, since YPHSM maintains a pH of 7.3 at 48-72 hours, suitable for high levels of stable protein expression. YPEM media becomes more acidic at 48 hours, leading to a possible degradation of the fusion protein by acid induced denaturation (47). It has been reported that some proteins only accumulate efficiently in rich media or minimal media buffered to near neutral pH (pH 6.0) (48). The media was therefore buffered at pH 6.0 to optimise the physiological environment for P45017α. SDS-PAGE analysis of these media revealed two induced protein bands corresponding to molecular masses in the 45 and 60 kDa range. It is possible that the latter was the expressed P45017α, as the calculated molecular mass of the FLAG-P45017α fusion protein is estimated to be ~59 kDa. This protein was, however, not detectable with the anti-FLAG monoclonal antibodies. The inability of the anti-FLAG monoclonal antibodies to detect the heterologously expressed P45017α could be ascribed to three factors. Firstly, the concentration of the heterologously expressed P45017α in the media could be too low for immunological detection. However, this was unlikely since the media was concentrated prior to immunological analysis. Secondly, the sensitivity of the detection method could have been too low. The method relies on a colorimetric detection system with 4-chloronaphtol as a substrate which is less sensitive than a chemiluminescent substrate such as luminol (49). Thirdly, the hydrophilic FLAG octapeptide could have been shielded by the large hydrophobic P45017α, hindering the binding of the FLAG antibody. Western blot analysis of the expression media using a nickel-HRP probe which recognises the five histidine tag on the carboxyl-terminal of FLAG-CYP17 fusion protein was also unsuccessful (results not shown).

The possibility that CYP17 was not transcribed or that a premature splicing of mRNA could have occurred was also investigated. Northern blot analysis of mRNA from yeast

Stellenbosch University http://scholar.sun.ac.za

cells in the exponential phase, grown in the rich media with glycerol (3% v/v), indicated that the CYP17 sequence was successfully transcribed and that the mRNA was intact.

It could be speculated that the *ADH2* promoter was not efficiently switched on in cells grown on glycerol, resulting in low initiation of CYP17 translation, leading to low expression levels in the medium. The *ADH2* promoter is one of many yeast genes of which expression is regulated by glucose repression. The transcription of *ADH2* is undetectable when yeast is grown on glucose and is derepressed to a level representing ~1% of soluble cellular protein when yeast is grown on a non-fermentable carbon source, e.g. glycerol (50). It has been shown that certain yeast strains harbouring the *ADH2* promoter are induced more efficiently with alternate non-fermentable carbon sources such as ethanol, raffinose or galactose (48). The physiological conditions and media were optimised by changing the carbon and non-fermentable carbon sources of either the expression or of the growth media in an effort to increase mRNA levels (different promoter induction). Dot blot analyses of CYP17 transcripts from cells grown in complex expression media supplemented with 2% raffinose and 2% ethanol showed considerably higher levels of mRNA than cells using glycerol. The highest level of CYP17 transcription was at 24 hours which can be ascribed to the strong facilitation of *ADH2* promoter induction.

Romanos *et al*. suggested that yeast harbouring a vector with an *ADH2* promoter need to be grown on plates containing excess glucose (8% w/v) in order to maintain repression (48). The same expression media, containing 2% raffinose and 2% ethanol, was therefore used in this investigation and, in addition, comparative analyses were carried out in cells grown on 8% glucose instead of 2% glucose, prior to induction. Dot blot analysis showed that the promoter is switched on much sooner as indicated by maximum transcription levels detected at 18 hours. Although cells grown on 3% glycerol and 0.8% galactose showed higher transcription levels than cells grown in original medium, 3% glycerol and 1% glucose, transcription levels were lower than that detected in cells grown on ethanol and raffinose.

Yeast cells harbouring a multi-copy episomal plasmid, YEp, tend to lose their plasmid during non-selective growth on complex medium, resulting in low transcription levels

Stellenbosch University http://scholar.sun.ac.za

(51). Cells were therefore also grown in synthetic selective medium supplemented with 0.8% galactose. No increase in expression levels were seen with dot blot analyses suggesting that plasmid stability was not the cause of low transcription levels. It was concluded from the dot blot analyses that the cells grown on ethanol and raffinose induced the *ADH2* promoter most efficiently and, in addition, Northern blot analyses confirmed the formation of full-length CYP17 transcripts of 2.2 kb.

It is likely that the BJ3505 yeast strain is not suitable for high levels of heterologous protein expression when expression is under the control of the *ADH2* promoter. It is commonly known that some strains are incompatible with certain promoter constructs and Plüddemann *et al*. observed low levels of transcription of a hepatitis B gene using the *ADH2* promoter with the BJ3505 strain. It was shown that by simply changing the yeast strain (Y294) mRNA levels were increased significantly (52)

Despite being able to increase the transcriptional level of CYP17 by optimising the media, immunodetection identified no expressed fusion protein in the media. This would suggest that the fusion protein still remained intracellular, either located in the periplasmic space or in the endoplasmic reticulum (ER).

S. cerevisiae has often been used as a host for the production and secretion of large heterologous proteins (> 250 kDa) into the medium (53). Many expressed proteins were, however, not able to cross the yeast cell wall and were therefore found in the periplasmic space. Molecular mass, shape and electric charge are characteristics that appear to affect the movement of the expressed proteins through the cell wall (54). The possibility that the expressed human P45017α was retained by the cell wall, due to relative high molecular mass and hydrophobic nature could however not be verified as the immunoblot of the fractions isolated from the periplasmic space fractions did not detect FLAG-P45017α fusion protein.

Unsuccessful secretion of the FLAG-P45017α fusion protein could also be due to failure of translocation across the ER membrane. The protein, with its hydrophobic membrane-binding domain, is likely to bind to the ER preventing it to be transported through the secretory pathway. Sakaguchi *et al*. provided evidence that a short amino terminal of

Stellenbosch University http://scholar.sun.ac.za

microsomal P450 protein serves not only as a signal for the membrane insertion but also as a "stop-transfer" sequence (55). The intracellular fraction was subsequently investigated and FLAG-P45017α fusion protein was not detected. In summary, no immunological detectable FLAG-P45017α protein was present in the intracellular or in the periplasmic fractions and it was therefore concluded that the secretion of heterologously expressed human cytochrome P45017α into the medium was unsuccessful.

A feasible explanation for the absence of heterologously expressed human P45017α was sought at the translational level — the encoded mRNA was not translated efficiently. CYP17 cDNA was correctly inserted in reading frame with the initiation codon of the MFα-1 secretory secretion gene. This AUG codon is the first AUG codon encountered by the scanning ribosomal complex. Investigation of the 5' untranslated region of the mRNA did not indicate any translational barriers such as of G residue repeats or the formation of secondary structures. Both can lead to a significant reduction in translation efficiency (56, 57). The use of non-preferred codons could also have contributed to the efficiency of translation of a specific mRNA in *S. cerevisiae* (58). Furthermore inadequate or incorrect folding of the expressed protein in the ER could result in the absence of the heterologously expressed fusion protein. Yeast proteins assisting in folding and disulphide bond formation differ from their counterparts in higher eukaryotes (e.g. mammalian cells). Malfolding may result in the retention of expressed fusion proteins in the ER and subsequent degradation (59).

The heterologous production of significant levels of fusion proteins in a yeast expression system had major limitations and was clearly unsuitable. Firstly, yeast cells that harbour episomal vectors, such as YEpFLAG, need to be grown under constant selective pressure in order to remain stable and ensure good expression levels. These conditions result in slow growth characterised by low yields. Secondly, glucose-repressible promoters, such as *ADH2,* are not suited for industrial fermentation processes since maintaining controlled glucose-repression under conditions of glucose limitations, required for high cell density, remains a formidable challenge (60). Thirdly, the YEpFLAG plasmid is an autonomously replicating vector. It has been shown that secreted proteins are produced at higher levels when the transcriptional unit is integrated

Stellenbosch University http://scholar.sun.ac.za

in the yeast chromosomes as apposed to an autonomously replicating vector (61). The former will provide greater vector stability. Fourthly, it is also possible that the ADH2 promoter and the yeast strain used in the expression of P45017α as a fusion protein are not compatible. Finally, the supplementation of the media with raffinose or galactose in an attempt to increase protein production would be very costly and is not commercially viable. Taking into consideration all of the aforementioned factors, a more efficient expression system for the heterologous expression and production of human P45017α was subsequently investigated.

Stellenbosch University http://scholar.sun.ac.za

REFERENCES

1 Zuber, M. X., Simpson, E. R. and Waterman, M. R. (1986) *Science* **234**, 1258-1261.

2 Bradshaw, K. D., Waterman, M. R., Couch, R. T., Simpson, E. R. and Zuber, M. X. (1987) *Mol. Endocrinol.* **1**, 348-354.

3 Zuber, M. X., Mason, J. I., Simpson, E. R. and Waterman, M. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 699-703.

4 Yanase, T., Kagimoto, M., Suzuki, S., Hashiba, K., Simpson, E. R. and Waterman, M. R. (1989) *J. Biol. Chem.* **264**, 8076-8082.

5 Kitamura, M., Buczko, E. and Dufau, M. L. (1991) *Mol. Endocrinol.* **5**, 1373-1380.

6 Clark, B. J. and Waterman, M. R. (1991) *J. Biol. Chem.* **266**, 5898-5904.

7 Trant. J. M. (1995) *J. Exp. Zool.* **272**, 25-33.

8 Barnes, H. J., Jenkins, C. M. and Waterman, M. R. (1994) *Arch. Biochem. Biophys*. **315**, 489-494.

9 Sakaki, T., Shibata, M., Yabusaki, Y., Murakami, H. and Ohkawa, H. (1989) *DNA Cell Biol.* **8**, 409-418.

10 Shibata, M., Sakaki, T., Yabusaki, Y., Murakami, H. and Ohkawa, H. (1990) *DNA Cell. Biol.* **9**, 27-36.

11 Swart, A. C., Swart, P., Roux, S. P.,van der Merwe, K. J., Pretorius, I. S. and Steyn, A. J. C. (1995) *Endocr. Res.* **21**, 289-295.

12 Auchus, R. J., Lee, T. C. and Miller, W. L. (1998) *J. Biol. Chem.* **273**, 3158-3165.

13 Barnes, H. J., Arlotto, M. P. and Waterman, M. R. (1991*) Proc. Natl. Acad. Sci. USA* **88**, 5597-5601.

14 Imai, T., Globerman, H., Gertner, J., Kagawa, N. and Waterman M. R. (1993) *J. Biol. Chem*. **268**, 17317-17375.

15 Trant, J.M. (1996) *Gen. Comp. Endocr.* **102**, 173-182.

16 Owaki, A., Takamasa, A., Yamazaki, T and Kominami, S. (2002) *J. Steroid Biochem.* **81**, 255-262.

17 Waterman, M. R. (1993) *Biochem. Soc. T.* **21**, 1081-1085.

18 Gonzalez, F. J. and Korzekwa, K. R. (1995) *Annu. Rev. Pharmacol.* **35**, 369-390.

19 Sagara, Y., Barnes, H. J. & Waterman, M. R. (1993) *Arch. Biochem. Biophys.* **304**, 272-278.

20 Pritchard, M. P., Ossetian, R., Li, D. N., Henderson, J., Burchell, B., Wolf, C. R. and Friedberg, T. (1997) *Arch. Biochem. Biophys.* **345**, 342-354.

21 Larson, J. R., Coon, M. J. and Porter, T. D. (1991) *J. Biol. Chem.* **266**, 7321-7324.

22 Iwata, H., Fujita, K., Kushida, H., Suzuki, A., Konno, Y., Nakamura, K., Fujino, A. and Kamataki, T. (1998) *Biochem. Pharmacol.* **55**, 1315-1325.

23 Smith, R. A., Duncan, M. J. and Moir, D. T. (1985) *Science* **229**, 1219-1224.

24 Brake, A. J. (1990) *Method. Enzymol.* **185**, 408-421.

25 De Smit, M. H. and van Duin, J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 7668-7672.

26. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

27. Becker, D. M. and Guarente, L. (1991) *Method. Enzymol.* **194**, 182-187.

28. Hoffman, C. S. and Winston, F. (1987) *Gene* **57**, 267-272.

29. Hoffman, C. S. (1997) *Curr. Protocol. Molec. Biol.*, 13.12.2.

30 La Grange, D. C., Claeyssens, M., Pretorius, I. S. and Van Zyl, W. H. (2000) Appl. Microbiol. Biot. **54**, 195-200

31.. Laemmli, U. K. (1970) *Nature* **227**, 680-685.

32. Oeda, K., Sakaki, T. and Ohkawa, H. (1985) *DNA Cell Biol.* **4,** 167-175.

33. Peyronneau, M.-A., Renaud, J.-P., Jaouen, M., Urban, P. and Cullin, C. (1993) *Eur. J. Biochem.* **218**, 355-361.

34. Renaud, J.-P., Cullin, C., Pompon, D., Beaune, P. and Mansuy, D. (1990) *Eur. J. Biochem.* **194**, 889-896.

35. Shimizu, T., Hirano, K., Takahashi, M., Hatano, M. and Fujii-Kuriyama, Y. (1988) *Biochemistry* **27**, 4138-4141.

36. Ellis, S. W., Ching, M. S., Watson, P. E., Henderson, C. J., Simula, A. P., Lennard, M. S., Tucker, G. T. and Woods, H. F. (1992) *Biochem. Pharmacol.* **44**, 617-620.

37. Goldstein, J. A., Faletto, M. B., Romkes-Sparks, M., Sullivan, T. and Kitareewan, S. (1994) *Biochemistry* **33**, 1743-1752.

38. Sakaki, T., Kominami, S., Takemori, S., Ohkawa, H., Akiyoshi-Shibata, M. and Yabusaki, Y. (1994) *Biochemistry* **33**, 4933-4939.

39. Pompon, D., Louerat, B., Bronine, A. and Urban, P. (1996) *Method. Enzymol.* **272**, 51-63.

40. Ohgiya, S., Kii, T., Goda, T., Hoshino, T., Hamabuchi, K., Yokota, H., Yuasa, A. and Ishizaki, K. (1997) *Biotechnol. Lett.* **19**, 437-441.

41. Guengerich, F. P., Brian, W. R., Sari, M.-A. and Ross, J. T. (1991) *Method. Enzymol.* **206**, 130-145.

42. Gelissen, G. (2000) *Appl. Microbiol. Biotech.* **54**, 741-750.

Stellenbosch University http://scholar.sun.ac.za

43. Lin Cereghino, G. P., Lin Cereghino, J., Ilgen, C. and Cregg, J. M. (2002) *Curr. Opin. Biotech.* **13**, 329-332.

44. Giga-Hama, Y. and Kumagai, H. (1999) *Biotechnol. Appl. Bioc.* **30**, 235- 244.

45. Bureik, M., Schiffler, B., Hiraoka, Y., Vogel, F. and Bernhardt, R. (2002) *Biochemistry* **41**, 2311-2321.

46. Zhang, W., Hywood Potter. K. J. H., Plantz, B. A., Schlegel, V. L., Smith, L. A. and Meagher, M. M. (2003) *J. Ind. Microbiol. Biotechnol.* **30**, 210-215.

47. Yeast N- Terminal FLAG Expression System manual, Eastman Kodak Companu, 8/94.

48. Romanos, M. A, Scorer, C. A and Clare, J. J. (1995) *DNA Cloning 2* (D. M. Glover and B. D. Hames, ed.) 2nd Ed, Oxford University Press Inc., New York.

49. Colyer, J. (1999) In: "*Protein Expression: A practical approach*" (S. J. Higgins and B. D. Hames, eds.) pp.225-265. Oxford University Press Inc., New York.

50. Price, V. L., Taylor, W. E., Clevenger, W., Worthington, M. and Young, E. T. (1990) *Method. Enzymol.* **185**, 308-318.

51. Tuite, M. J., Clare, J. J. and Romanos, M. A. (1999) In: "*Protein Expression: A practical approach*" (S. J. Higgins and B. D. Hames, eds.) pp.225-265. Oxford University Press Inc., New York.

52 Plüddemann, A. (1997) M. Sc. thesis, University of Stellenboch.

53. Jabbar, M. A. and Nayak, D. P. (1987) *Mol. Cell. Biol.* **7**, 1476-1485.

54. De Nobel, J. G. and Barnett, J. A. (1991) *Yeast* **7**, 313-323.

55. Sakaguchi, M., Mihara, K and Sato, R. (1987) *EMBO J.* **6**, 2425-2431.

56. Cigan, A. M. and Donahue, T. D. (1987) *Gene*, **59**, 1-18.

57. Vega-Laso, M. R., Zhu, D., Sagliocco, F., Brown, A. J. P., Tuite, M. F. and McCarthy, J. E. G. (1993) *J. Biol. Chem.* **268**, 6453-6462.

58. Buckholz, R.G. and Gleeson, M.A.G. (1991) *Bio/Technology* **9**, 1067-1072.

59. Moir, D.T. (1989) In: P.J Barr, A.J. Brake and P. Valenzuela, P. (Eds), *Yeast Genetic Engineering*, pp. 215-231. Butterworths, Boston.

60. Moir, D.T. and Davidow, L.S. (1991) *Methods in Enzymology* **194**, 491-519.

61. Smith, R.A., Duncan, M.J. and Moir, D.T. (1985) *Science* **229**, 1219-1224.

Stellenbosch University http://scholar.sun.ac.za

PART III

CLONING AND EXTRACELLULAR EXPRESSION OF THE FULL-LENGTH HUMAN CYTOCHROMES P45017α, AROMATASE, b_5 AND TRUNCATED HUMAN CYTOCHROME P45017α IN *PICHIA PASTORIS*

Stellenbosch University http://scholar.sun.ac.za

INTRODUCTION **132**

CHAPTER 1 : PLASMID PREPARATIONS, TRANSFORMATION AND SCREENING OF *P. PASTORIS* **137**

1.1 Cloning and construction of the full-length human CYP17 into pPIC9K expression vector and integration in *P. pastoris* **137**

1.1.1 **Materials and Methods** **137**

1.1.1.1 Strains and media 137

1.1.1.2 Subcloning of the CYP17 cDNA 138

1.1.1.2.1 Polymerase chain reaction (PCR) amplification 138

1.1.1.2.2 Restriction digestion 139

1.1.1.2.3 Dephosphorylation 139

1.1.1.2.4 Ligation 140

1.1.1.2.5 Analyses of correct clones 141

1.1.1.3 Construction of the pPIC9K-CYP17 expression vector 141

1.1.1.4 Amplification of pPIC9K-CYP17 in *E. coli.* 143

1.1.1.4.1 Transformation 143

1.1.1.4.2 Isolation 143

1.1.1.4.3 DNA sequencing analysis of E. coli transformants 143

1.1.1.5 Screening of GS115/pPIC9K-CYP17 and KM71/pPIC9K-CYP17 transformants for integration 144

1.1.2 **Results** **145**

1.1.2.1 Subcloning of the CYP17 cDNA 145

1.1.2.2 Construction of the pPIC9K-CYP17 expression vector 147

1.1.2.3 Screening of GS115/pPIC9K-CYP17 and KM71/pPIC9K-CYP17 transformants for integration 148

1.2 Cloning and construction of the truncated human CYP17 (truncCYP17) into the pPIC9K expression vector **152**

Stellenbosch University http://scholar.sun.ac.za

_1.2.1_ Materials and methods **152**

1.2.1.1 Subcloning of the truncCYP17 cDNA 152

1.2.1.2 Construction of the pPIC9K-truncCYP17 expression vector 154

1.2.1.3 Screening of GS115/pPIC9K-truncCYP17 and KM71/pPIC9K-truncCYP17 transformants for integration 156

_1.2.2_ Results **156**

1.2.2.1 Subcloning of the truncCYP17 cDNA 156

1.2.2.2 Construction of the pPIC9K-truncCYP17 expression vector 157

1.2.2.3 Screening of GS115/pPIC9K-truncCYP17 transformants for integration 157

1.3 Cloning and construction of the full-length human b_5 gene (CYB5) into the pPIC9K expression vector **159**

_1.3.1_ Materials and Methods **159**

1.3.1.1 Subcloning of the CYB5 cDNA 159

1.3.1.2 Construction of the pPIC9K-CYB5 expression vector 161

1.3.1.3 Screening of GS115/pPIC9K-CYB5 and KM71/pPIC9K-CYB5 transformants for integration 163

_1.3.2_ Results **163**

1.3.2.1 Subcloning of the CYB5 cDNA 163

1.3.2.2 Construction of the pPIC9K-CYB5 expression vector 164

1.3.2.3 Screening of GS115/pPIC9K-CYB5 and KM71/pPIC9K-CYB5 transformants for integration 166

1.4 Cloning and construction of full-length human CYP19 into the pPIC9K expression vector **168**

_1.4.1_ Materials and Methods **168**

1.4.1.1 Subcloning of the CYP19 cDNA 168

1.4.1.2 Construction of the pPIC9K-CYP19 expression vector 171

1.4.1.3 Screening of GS115/pPIC9K-CYP19 and KM71/pPIC9K-CYP19 transformants for integration 172

Stellenbosch University http://scholar.sun.ac.za

***1.4.2* Results** **173**

1.4.2.1 Subcloning of the CYP19 cDNA 173

1.4.2.2 Construction of the pPIC9K-CYP19 expression vector 174

1.4.2.3 Screening of GS115/pPIC9K-CYP19 and KM71/pPIC9K-CYP19 transformants for integration 174

CHAPTER 2 : CULTURE OF *P. PASTORIS* EXPRESSING DIFFERENT PROTEINS **177**

2.1 Growth and expression of P45017α, aromatase and b_5 in *P. pastoris* **177**

***2.1.1* Materials and Methods** **177**

2.1.1.1 Growth 177

2.1.1.2 Screening of yeasts for secretory expression of proteins 177

2.1.1.3 Screening for intracellular expression of the proteins 178

2.1.1.4 SDS-PAGE and Western blot analyses 178

***2.1.2* Results** **179**

2.1.2.1 Growth 179

2.1.2.2 Expression of full-length and truncated P45017α by GS115 cells 181

2.1.2.3 Expression of full-length P45017α by *P. pastoris* KM71 183

2.1.2.4 Expression of b_5 by *P. pastoris* GS115 183

2.1.2.5 Expression of b_5 by *P. pastoris* KM71 184

2.1.2.6 Expression of aromatase by *P. pastoris* GS115 185

2.1.2.7 Expression of aromatase by *P. pastoris* KM71 186

2.1.2.8 Control strains 187

CHAPTER 3 : DISCUSSION **188**

REFERENCES **193**

Stellenbosch University http://scholar.sun.ac.za

INTRODUCTION

In human adrenal and gonadal steroidogenesis, two key hemeproteins, P45017α and aromatase are involved in sex steroid biosynthesis. P45017α converts C_{21} steroids to C_{19} sex steroids in two different reactions. Firstly, pregnenolone and progesterone are converted to 17-hydroxylated intermediates (17-hydroxypregnenolone and 17-hydroxyprogesterone) catalysed by a 17α-hydroxylase reaction. Subsequently only 17-hydroxypregnenolone is further catalysed by a 17,20-lyase reaction to DHEA (1). The conversion of the second hydroxylated intermediate, 17-hydroxyprogesterone to androstenedione is found to be insignificant in humans and rodents (2, 3, 4). DHEA is catalysed to androstenedione by 3β-HSD, which is converted to oestadiol by aromatase (5).

The 17α-hydroxylation reaction that directs pregnenolone and progesterone to the glucocorticoids has been well characterised, but the subsequent 17,20-lyase reaction that converts 17-hydroxypregnenolone to the sex steroid precursor, DHEA, is still under investigation (6). An understanding of the regulation of the 17,20-lyase activity very important, as it is the sole pathway to the biosynthesis of sex steroids thereby regulating the onset of DHEA formation during adrenarche and the physiological development of polycystic ovary syndrome (PCOS) (7). Furthermore, the regulation of the 17,20-lyase activity seems to be under the control of the rate of flow of electrons to the enzyme and it has been suggested that b_5 may function as an alternative source of electrons specific for the 17,20-lyase reaction but not the 17α-hydroxylase reaction (8, 9). The conversion of the androgen, androstenedione, to oestrogens, a reaction known as aromatisation, is a key regulatory step in oestrogen biosynthesis. Aromatase has been the focus of much interest since the correlation of oestrogen levels to human disease conditions, such as breast and endometrial cancer is well established (10, 11, 12). A better understanding of the relationship of structure to function in P45017α, b_5 and aromatase, will enable researchers to gain a better understanding in the involvement of these complex enzymes in their respective reactions. The availability of purified aromatase will permit the development of specific inhibitors which could be used in the prevention of human breast, endometrial and ovarian cancers (13, 14, 15, 16, 17). In addition, poly- and monoclonal antibodies against these enzymes would assist in diagnostic applications

Stellenbosch University http://scholar.sun.ac.za

such as immunohistochemistry and specific targeting (18). Monoclonal antibodies against aromatase could also be used to identify the exact location of intratumoral aromatisation in clinical samples of breast cancer.

Several research groups have attempted to isolate and purify aromatase from human placenta, encountering the following obstacles. Firstly, the specific content of their preparations was low (19, 20). Secondly, only a partially purified human placental microsomal aromatase was obtained (21, 22). Thirdly, although a homogenous aromatase protein was produced, it was either unstable (23) or the recovery after substrate affinity chromatography purification was low (24). Kellis and Vickery (25) and Yoshida and Osawa (26) were able to obtain a pure, catalytically active aromatase in a stable form from human placenta. The isolation of P45017α from tissue was unsuccessful, since the concentrations were very low.

To circumvent the experienced difficulties of isolating these eukaryotic cytochromes P450 directly from their specific tissues, scientists relied on molecular biology to synthesise specific forms of cytochromes P450 in various heterologous expression systems. *E. coli* (27,), *S. cerevisiae* (28, 29, 4), COS 1 (30) and baculoviral cells (31, 32, 33,) have been used as hosts for the expression of human P45017α, b_5 and aromatase. However, various shortcomings which include inactive enzyme, low concentrations and improper folding limited their usefulness as systems for large-scale production and therefore rendered them unsuitable for antibody production.

Extracellular expression systems were considered to overcome the obstacles of intracellular expression. Secreting expressed cytochromes P450 into the culture medium via the hosts' secretory pathway has the following advantages: (i) secreted cytochromes P450 are more readily available in the essentially protein-free culture medium, leading to minimal purification steps from the expression medium, (ii) the usage of harsh denaturants and detergents to facilitate breakage of cells and creating an unsuitable oxidising environment might be avoided, (iii) proper protein folding and translational modifications of the P450 enzyme will be increased and protection from cytoplasmic proteases will take place (34) and (iv) the cytoplasmically expressed protein products may exert a toxic or unstable effect in the host, which can be avoided when they are

Stellenbosch University http://scholar.sun.ac.za

secreted into the medium (34). High yields of stable, intact cytochromes P450 may therefore be obtained.

These requirements for high-level extracellular expression of recombinant proteins are met by the methylotrophic yeast *Pichia pastoris* (35, 36). Recently, *P. pastoris* has gained widespread attention as an extracellular expression system which secreted many different types of proteins, including membrane proteins (37, 38, 39) into the medium. One of the reasons why *P. pastoris* could be successful in secreting proteins extracellulary is that its ultrastructural organisation differs from that observed in *S. cerevisiae*. *P. pastoris* has a similar transitional ER-Golgi system to that of vertebrate cells (40) and the secretory vesicles of *P. pastoris* are three times larger than that of *S. cerevisiae* (41). Furthermore, *P. pastoris* offers many additional advantages — (i) the heterologous genes are integrated into the yeast genome downstream of an alcohol oxidase 1 (*AOX1*) gene promoter, which ensures stable integrations for many generations and tightly regulated methanol-induced expression at very high levels, (ii) maintenance in large-volume fermenter cultures and at cell densities in excess of 100-500 grams/L (42) and (iii) *P. pastoris* secretes very low levels of native proteins. Therefore, the isolation and purification of the expressed protein from the cellular proteins is easier and more effective in the *P. pastoris* system than in *S. cerevisiae*.

There are three major criteria in secretion of expressed proteins that may interfere or limit the extracellular expression of cytochromes P450. Hydrophobicity, protein size and glycosylation, which are the focus of this study.

A hydrophobic sequence in the protein can bind to the ER thereby preventing the movement to the outside of the cell. Eukaryotic microsomal cytochrome P450 proteins are integral membrane proteins of the ER and are targeted to the ER via the signal recognition particle pathway (43). The targeting and the insertion of these cytochromes P450 are defined by a hydrophobic sequence (signal anchor sequence) which is usually situated in the amino-terminal of the protein (44, 45). Several cytochromes P450 lacking the hydrophobic amino-terminal region have been studied in various laboratories and in most cases the heterologously expressed, modified proteins retained their catalytic activity and membrane-binding capability, as summarised by Pernecky and Coon. (46).

Stellenbosch University http://scholar.sun.ac.za

However, studies have shown that the removal of portions of the putative membrane-anchor domain from microsomal cytochromes P450 rendered these proteins incapable of membrane insertion and were primarily located in the soluble fraction in *E. coli* (47, 48, 49, 50) and *S. cerevisiae* (51).

We therefore investigated the effect of the hydrophobic domain of P45017α on the solubility of the protein by comparing the full-length P45017α with a truncated form (deletion of first 18 amino acids) by expressing the two forms of P45017α extracellularly. By deleting the hydrophobic membrane-binding domain, the chances of the expressed truncated P45017α binding to the membranes of the secretory organelles, while being transported through the *Pichias'* secretory pathway to the outside of the cells, could be minimised and higher expression levels could be obtained in the medium.

Protein size has also been documented to influence extracellular expression in yeast (52). It is not a problem for large secreted proteins to be transported through the yeast's secretory compartments. The crossing of the yeast cell wall of expressed proteins does, however, present an obstacle. In earlier studies it was believed that secreted proteins in *S. cerevisiae*, in excess of 30 kDa, are retained within the cells or the periplasmic space. Recent findings, however, show molecules with relative masses as great as 400000 kDa often diffuse through the yeast cell wall quite easily (52). The reason for this is that although the average pore sizes of the cell wall are small, the permeability of the cell walls vary, depending on the yeast strain and its physiological conditions, i.e. growth conditions. In the exponential growth phase very large proteins are able to passage over the cell wall compared to non-growing yeasts which retain the large proteins. Furthermore, the biochemical make-up of the heterologous protein secreted, such as molecular mass, shape and electric charge appears to affect movement through the cell wall. The b_5, which is a fairly small protein (15199 Da), might be secreted more readily than the bigger sized P45017α (~57 kDa) or aromatase (~58 kDa)

In *P. pastoris* the problem of glycosylation is far less than in *S. cerevisiae* (53). The reason being that instead of an average oligosaccharide chain length of 40 mannose residues from *S. cerevisiae*, glycoproteins from *P. pastoris* synthesise an average chain length of 8-12 mannose residues. Therefore *P. pastoris*, unlike the wild-type *S.*

Stellenbosch University http://scholar.sun.ac.za

cerevisiae, does not hyperglycosylate (outer chain glycosylation) secreted glycoproteins. It appears that the glycosylation process of *P. pastoris* is closer to the mammalian high-mannose type than to that of *S. cerevisiae*. Shimozawa *et al*. found that aromatase purified from human placenta microsomes is glycosylated (54) and Gartner *et al.* also produced a glycosylated aromatase protein which was heterologously expressed in a baculovirus expression system (33).

In this study we investigated these abovementioned three criteria, i.e. hydrophobicity, size and glycosylation, by heterologously expressing a full-length P45017α, a truncated P45017α, b_5 and aromatase, which would be secreted into the media by two different *P. pastoris* yeast strains. The cDNA of the different steroidogenic enzymes were cloned in reading frame to the α-factor secretory signal sequence which directs the translated proteins to the medium under the control of a MFα-factor secretory signal. The fusion proteins with histidine affinity tags at the carboxyl end of the protein allowed for immunological detection of the expressed proteins. The culture media were screened for the presence of secreted cytochromes P450. In addition the cytoplasmic fractions were analysed for the presence of expressed cytochromes P450. Comparisons between GS115 and KM71 cells were also investigated.

This is the first report of an attempt to express human P450 enzymes extracellularly in a yeast system.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 1 : PLASMID PREPARATIONS, TRANSFORMATION AND SCREENING OF *P. PASTORIS*

1.1 Cloning and construction of the full-length human CYP17 into pPIC9K expression vector and integration in *P. pastoris*

1.1.1 Materials and Methods

1.1.1.1 Strains and media

The *E. coli* strain JM109 (Promega) was used in all DNA manipulations. For routine growth in *E. coli*, sterile Luria-Bertani (LB) medium (1% Bacto-Tryptone, 0.5% Bacto Yeast Extract, 1% NaCl, pH 7.0) was used. All media components were purchased from Difco. Selection of bacteria transformants was carried out on LB-agar (15g/L) containing ampicillin (100 μg/ml). For blue /white selection X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside), 20 mg/ml, was added to LB-agar.

P. pastoris strains GS115 (*his4*) and KM71 (*arg4 his4 aox1:: ARG*), used for the expression of human cytochrome P45017α, were grown, transformed and analysed according to the manufacturer's (Invitrogen) instructions. For the routine growth of yeast, YPD media (1% yeast extract, 2% peptone, and 2% glucose) was used. Selection of yeast transformants was carried out in minimal dextrose (MD) medium containing per liter: 13,4 g yeast nitrogen base without amino acids, 400 μg biotin and 20 g dextrose. Minimal methanol (MM) medium contained 13.4 g yeast nitrogen base without amino acids, 400 μg biotin and 5 ml methanol per liter. Regeneration dextrose base (RDB) plates contained 1 M sorbitol, 2% dextrose, 1.34% yeast nitrogen base without amino acids, 4×10^{-5}% biotin, L-glutamic acid, L-methionine, L-lysine, L-leucine, L-isoleucine each at 0.005% and 2% agar. For heterologous expression *P. pastoris* was grown in buffered minimal glycerol (BMGY) medium (0.1 M potassium phosphate pH 6.0, 1.34% w/v yeast nitrogen base without amino acids, 4×10^{-5}% biotin, 1% v/v glycerol) or in

Stellenbosch University http://scholar.sun.ac.za

buffered minimal methanol medium without glycerol, containing 0.5% v/v methanol (BMMY).

1.1.1.2 Subcloning of the CYP17 cDNA

All DNA manipulations were carried out using standard procedures (55).

1.1.1.2.1 Polymerase chain reaction (PCR) amplification

The full-length CYP17 in the pFLAG-CYP17 construct, cloned in this laboratory, was used as the template to amplify a full-length CYP17 cDNA, depicted in Figure 1.1. The upstream primer STE1 (see Appendix B) and the downstream primer P104 (5'-CCGGAATTCTTA**TGATGGTGATGGTG**ATGGGTGCT-3') were used. The STEI and P104 primers introduced an *Eco*RI site (underlined) and the latter primer introduced a nucleotide sequence encoding six histidine residues (bold).

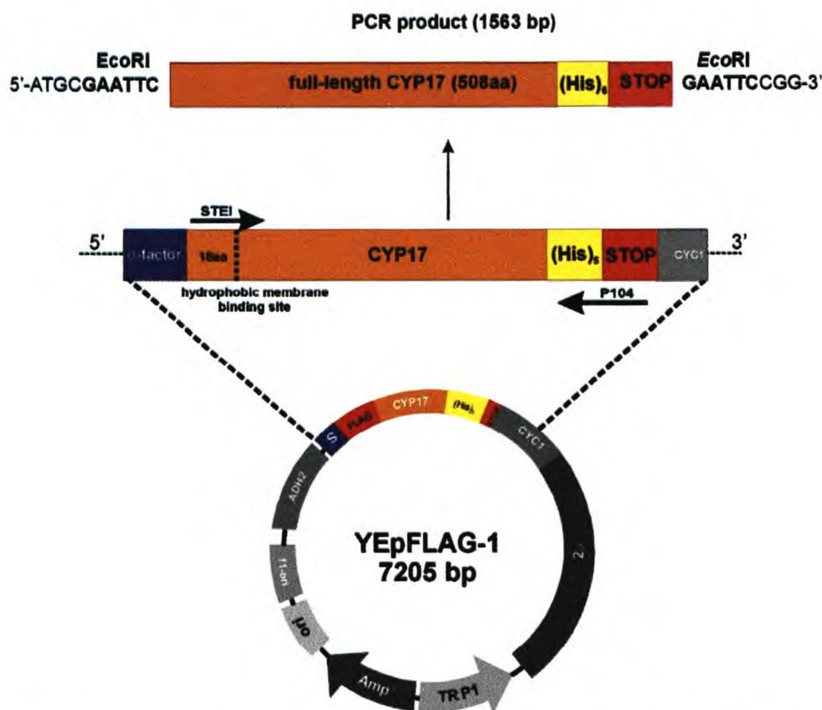

Figure 1.1 Schematic representation of the cloning strategy for the PCR amplification of CYP17 cDNA from the YEpFLAG-CYP17 vector. The upstream primer STE1 is complementary to the unmodified 5'-terminal and includes the start codon. The downstream primer P104 is complementary to the 3'-terminal, including the stop codon and nucleotide bases encoding six histidine residues. Both primers include a nucleotide sequence encoding an *Eco*RI site for subcloning into the pUC18 vector.

Each PCR amplification mixture (100 μl) contained pFLAG-CYP17 vector DNA (750 ng), 200 μM each of dGTP, dATP, dTTP and dCTP, 0.6 μM each of the sense and

Stellenbosch University http://scholar.sun.ac.za

antisense primers and 2.5 U proof-reading enzyme Pwo-polymerase (Roche) in Pwo buffer containing $MgSO_4$ (20 mM). The following procedure was carried out in a PCR-Sprint thermo cycler (Hybaid): (1) denaturation of the template at 94°C for 3 min; (2) 35 cycles of denaturing at 94°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 3 min; (3) and a final extension at 72°C for 10 min. The amplified DNA fragment was purified using an 0.8% low-melting agarose gel in 40 mM TAE buffer (Tris-HCl pH 8.0, glacial acetic acid and 5 mM EDTA). The DNA samples (1 μg), diluted in 0.2% loading buffer (0.1% Orange G (w/v), 20% Ficoll (w/v), 10 mM EDTA, pH 7.0), were electropheresed (EC Minicell® EC370M), at 55 V for 90 min, stained in ethidium bromide (10 μg/μl) and visualised under a UV lamp. The CYP17 fragment was excised and purified using the Wizard™ PCR Preps DNA Purification system (Promega), according to the manufacturer's instructions. The DNA was eluted with deionised water, 50 μl, and stored at –20°C.

1.1.1.2.2 Restriction digestion

The PCR product was digested with *Eco*RI yielding CYP17 cDNA fragment containing *Eco*RI compatible 5' and 3' termini (Figure 1.2), which was subsequently cloned into the pUC18 vector, linearised with *Eco*RI. Two separate restriction digestions, one containing amplified CYP17 DNA (2 μg) and the other pUC18 DNA (2μg)(Promega) were carried out at 37°C for 90 min, each in a final volume of 30 μl containing 5 μl restriction buffer B (supplied), acetylated BSA (5 μg) and *Eco*RI (10 U), (Promega). The digested DNA fragments were electrophoresed on an 0.8% low-melting agarose gel in TAE buffer (40 mM), and the digested CYP17 fragment and linearised vector were extracted and purified as described in 1.1.1.2.1.

1.1.1.2.3 Dephosphorylation

The linearised pUC18 vector was dephosphorylated prior to ligation to prevent religation without insert DNA in downstream reactions. Dephosphorylation mix (50 μl) containing DNA (1.8 μg), 5 μl reaction buffer (50 mM Tris-HCl, pH 9.3, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$ and 1 mM spermidine) and calf intestinal alkaline phosphatase (CIAP) (0.02 U) (Promega) was incubated at 37°C for 30 min. Another 0.02 U calf intestinal alkaline phosphatase was added to the mixture and incubated for 30 min. The dephosporylation

Stellenbosch University http://scholar.sun.ac.za

reaction was terminated as follows: 2 µl EDTA (0.5M) was added and the resultant reaction mixture was incubated at 65°C for 20 min. The DNA was purified from the mixture with the Magic DNA Clean-Up® kit (Promega), according to the manufacturer's instructions.

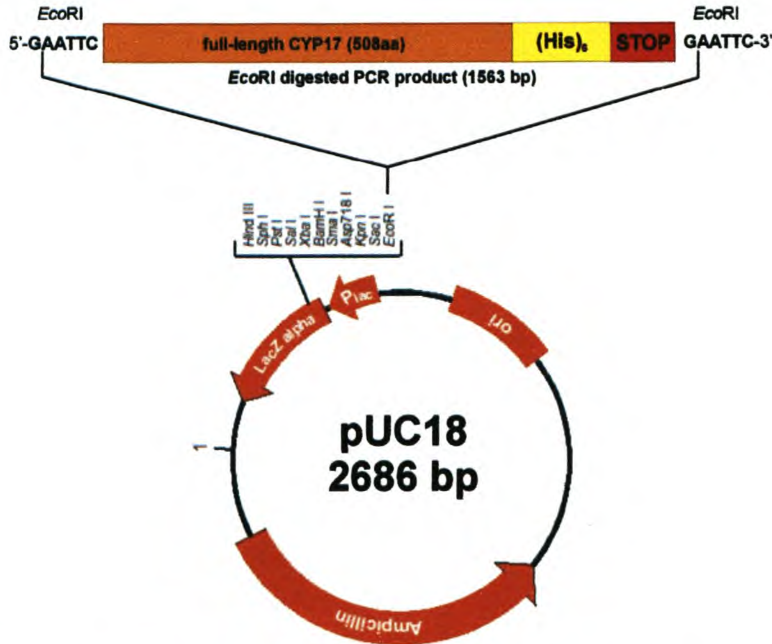

Figure 1.2 Schematic representation of the subcloning of the full-length CYP17 cDNA into pUC18 vector. The 3'-terminal of CYP17 contains a nucleotide sequence encoding the six histidine tag incorporated upstream of the stop codon.

1.1.1.2.4 Ligation

The purified CYP17 cDNA was subsequently ligated using 1:1 and 1:3 molar ratio of vector:insert DNA. For the 1:1 ratio DNA insert (112 ng) and vector (200 ng) was used in a 10 µl ligation reaction containing T4 ligase (1.5 U)(Promega), 0.5 µl ligase buffer (30 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 10 mM DTT and 1 mM ATP) and ATP (0.5 mM). The reaction was incubated at 16°C for 18 hours. Competent JM109 cells were transformed with the ligated vectors. In addition to ampicillin selection, the vector contained the lacZ gene, which allowed for blue/white colony selection.

The following ligation reactions were set up: A) dephosphorylated plasmid DNA (200 ng) without insert (negative control), B) dephosphorylated plasmid DNA (200 ng) with insert DNA (200 ng) carrying compatible phosphorylated termini and C) linearised

Stellenbosch University http://scholar.sun.ac.za

plasmid DNA (200 ng) not treated with CIAP. A positive control ligation and control transformations were also included.

1.1.1.2.5 Analyses of correct clones

Positive colonies were subsequently incubated overnight in LB medium (5 ml) at 37°C and the plasmid DNA was isolated using the Wizard™ *Plus* Minipreps DNA Purification System (Promega), according to the manufacturer's instructions. The purified recombinant plasmid DNA was digested with *Xba*I and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit (Macherey-Nagel) as described in the manual for high-copy number plasmids. The purified isolated pUC18-CYP17 plasmid was further analysed by PCR amplification.

The pUC18-CYP17 vector was analysed by PCR using the forward primer STE1 and reverse primer P104. Each PCR amplification mixture (50 µl) contained pUC18-CYP17 DNA (200 ng), 200 µM each of dATP, dTTP, dCTP and dGTP, 2.5 µM each of the STE1 and P104 primers and Supertherm DNA Taq polymerase (2.5 U) (Abgene) in TAPS (25 mM), pH 9.3, containing 50 mM KCl, 1 mM β-mercaptoethanol and 2.5 mM $MgCl_2$. The following procedure was carried out in a MiniCycler™ (MJ Research): (1) denaturation at 94°C for 3 min; (2) 35 cycles of denaturing at 94°C for 30 s, annealing at 60°C for 30 s and extension at 72°C for 3 min; (3) and a final extension at 72°C for 15 min. The PCR product was analysed by agarose gel electrophoresis (result not shown).

1.1.1.3 Construction of the pPIC9K-CYP17 expression vector

The cloned CYP17 in pUC18-CYP17 construct was confirmed by DNA sequence analyses. CYP17 was subsequently excised from the vector with *Eco*RI and ligated into pPIC9K vector (Invitrogen) previously linearised with *Eco*RI as shown in Figure 1.3.

Two separate restriction digestions, one containing pUC18-CYP17 DNA (2 µg) and the other containing pPIC9K DNA (2µg) were carried at 37°C for 90 min, each in a final volume of 30 µl containing 5.0 µl restriction buffer B (supplied), acetylated BSA (5.0

Stellenbosch University http://scholar.sun.ac.za

μg) and *Eco*RI (10 U). The digested CYP17 fragment and linearised pPIC9K vector were electrophoresed on an 0.8% low-melting agarose gel in TAE buffer (40 mM), extracted and purified using the Wizard™ PCR Preps DNA Purification system.

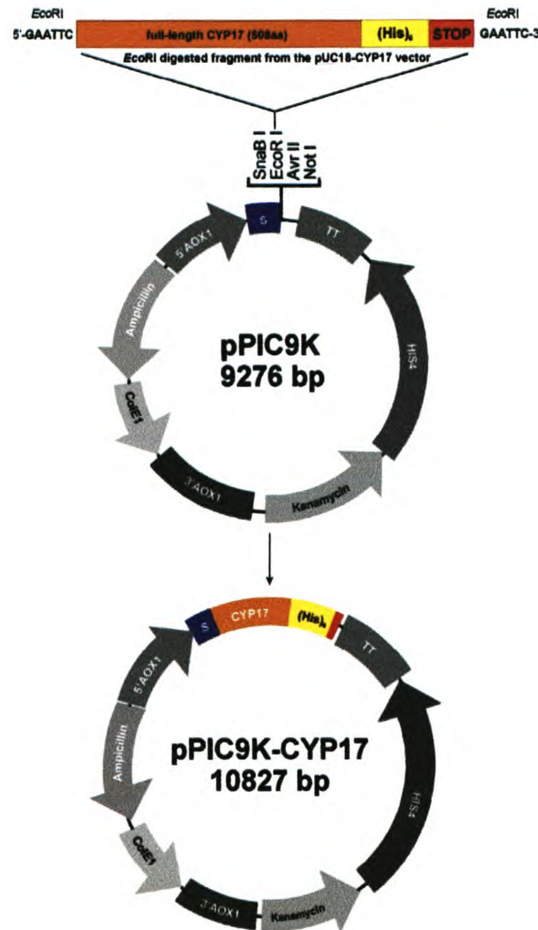

Figure 1.3 Schematic representation of the cloning strategy for the construction of the pPIC9K-CYP17 expression vector. The cloned CYP17 in the pUC18 vector was digested with *Eco*RI and inserted into the pPIC9K vector at the *Eco*RI site. The CYP17 was cloned in reading frame of the α-factor secretion signal and contains a nucleotide sequence encoding the six histidine tag.

The linearised pPIC9K DNA was dephosphorylated prior to ligation as described in section 1.1.1.2.3. The dephosphorylated linear vector was purified with the Magic DNA Clean-Up® kit (Promega), according to the manufacturer's instructions and analysed by agarose gel electrophoresis.

The purified digested CYP17 fragment (32.3 ng) and pPIC9K DNA (200 ng) were subsequently ligated as described in section 1.1.1.2.4. Competent JM109 cells were

Stellenbosch University http://scholar.sun.ac.za

transformed with the ligated vectors and selected on LB-agar plated containing ampicillin.

1.1.1.4 Amplification of pPIC9K-CYP17 in *E. coli.*

1.1.1.4.1 Transformation

Competent bacteria were prepared and the transformation was carried out as described in protocol 1 (55). Competent bacterial cell suspension (50 µl) was added to 5 µl of the ligation reactions, mixed gently and kept on ice for 30 min. The mixtures were subsequently heat shocked at 42°C for 90 s and incubated in ice for 90 s. After the addion of 800 µl SOC medium (37°C) (2% tryptone, 0.5% yeast extract and 0.05% NaCl, 250 mM KCl, pH 7.0, 10 mM $MgCl_2$ and 20 mM glucose), the cultures were incubated at 37°C for 45 min in a shaking incubator at 225 rpm. The cell suspension (100 and 150 µl) was plated out on LB-agar medium containing ampicillin (100 µg/ml) and incubated at 37°C overnight.

1.1.1.4.2 Isolation

Positive colonies were subsequently incubated overnight in LB medium (5 ml) at 37°C and the plasmid DNA was isolated as described in 1.1.1.2.5. The purified recombinant plasmid DNA was digested with *Xba*I and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit as described in section 1.1.1.2.5.

1.1.1.4.3 DNA sequencing analysis of E. coli transformants

The purified pPIC9K-CYP17 plasmid (100 ng/µl), together with its insertion sites was analysed by DNA sequencing using eight primers (1.1 pmol/µl) (Figure 1.4). The sense primers were STE1, 4011, 1496 and the antisense primers were 3989, 8982 and STE2. The upstream and downstream sequencing primers 5' α-factorP (5'-TACTATTGCCAGCATTGCTGC-3') and 3'AOX1P (5'-GCAAATGGCATTC TGACATCC-3') were also used for the sequencing reaction. Nucleotide sequences were

Stellenbosch University http://scholar.sun.ac.za

determined using the Bigdye™ Version 2 diterminator sequencing kit (model 373A ABI, Applied Biosystems, Foster City, CA).

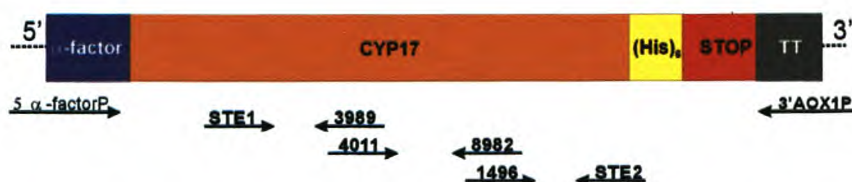

Figure 1.4 Schematic representation of the DNA sequencing strategy for the full-length human CYP17 in the pPIC9K-CYP17 expression vector. Primers α-factor and 3'*AOX 1* sequenced from the outside of the flanking regions of the CYP17 insert confirmed correct insertion. The remaining six primers sequenced the CYP17.

1.1.1.5 Screening of GS115/pPIC9K-CYP17 and KM71/pPIC9K-CYP17 transformants for integration

The cloned expression vectors, pPIC9K-CYP17 (10 μg) and the parent vector pPIC9K (10 μg), were linearised with *Sal*I and transformed into competent *P. pastoris* strains, GS115 Mut^+ (Methanol utilisation plus) and KM71 Mut^s (Methanol utilisation slow) using the spheroplast method as described by *Pichia* Spheroplast Module (Invitrogen). Controls (no DNA and parent vector without insert) were included to monitor contaminants and were used as background controls for expression. Before the linearised DNA was transformed into *P. pastoris* cells, spheroplasts were prepared by partially digesting the cell wall with Zymolyase (supplied with spheroplast kit) according to the manufacturer's instructions.

Transformants were plated on a selective medium lacking histidine, i.e. on minimal dextrose RDB plates, and the plates were incubated for 4 days at 30°C to select for His^+ recombinants. Transformants were then pooled and plated on YPD-G418 plates containing G418 (of 0.25 mg/ml). The level of G418 resistancy correlates with the number of kanamycin genes integrated into the *Pichia* genome. A single copy of pPIC9K plasmid integrated confers resistance to G418 to a level of 0.25 mg/ml. Screening for methanol utilisation, the single colonies (His^+ transformants) were picked up with a sterile toothpick and patched in a regular pattern on both MM plates and MD plates. The same procedure was used for the two control strains GS115/His^+ Mut^s

Stellenbosch University http://scholar.sun.ac.za

Albumin and GS115/His$^+$ Mut$^+$ β-gal. The plates were incubates at 30°C for 2 days and scored for Mut$^+$ and Mut^s phenotype by comparing the transformed colonies with the controls.

The presence of cDNA encoding the appropriate cytochrome genes integrated into the yeast chromosomes was subsequently confirmed by PCR amplification analyses. The protocol for direct PCR Screening was adapted from a method as described by Linder *et al.* (56). A pin-point-sized part of six single colonies from MD plates were selected, suspended in sterile water (30 µl) and lyticase (100 U) (crude, from *Arthrobacter luteus*; Sigma) added. The mixture was subsequently incubated at 30°C for 12 min. The digested cells were frozen in liquid nitrogen for 1 min and the lysate (5 µl) was used for the PCR amplification.

Each PCR amplification mixture (50 µl) contained of DNA/cell lysate (5 µl), 200 µM each of dGTP, dATP, dTTP and dCTP, 2.5 µM each of the 5'AOX1P and 3'AOX1P primers and of Taq-polymerase (2.5 U) in reaction buffer containing $MgCl_2$ (25 mM). The following procedure was carried out in a PCR-Sprint thermo cycler: (1) denaturation template at 94°C for 3 min; (2) 35 cycles of denaturing at 94°C for 1 min, annealing at 54°C for 1 min and extension at 72°C for 1 min; (3) and a final prolonged elongation time at 72°C for 7 min. The amplified DNA fragment was analysed by agarose gel electrophoreses as described in section 1.1.1.2.5.

Single colonies that contained the verified recombinants in GS115 or KM71 were grown in MD (10 ml) overnight at 30°C. The cells were isolated by centrifugation at 3000 × g for 5 min at room temperature and resuspended in fresh MD containing 15% glycerol. Cells were frozen in liquid nitrogen and stored at −80°C.

1.1.2 Results

1.1.2.1 **Subcloning of the CYP17 cDNA**

The full-length human CYP17, located in the pFLAG-CYP17 expression vector cloned in this laboratory, was used as the template to amplify a full-length CYP17 cDNA. The

Stellenbosch University http://scholar.sun.ac.za

predicted size of the amplified CYP17 cDNA product was ~1563 bp which included the nucleotides encoding the restriction enzymes. This was confirmed by 0.8% agarose gel electrophoresis (Figure 1.5 lanes 3 and 4).

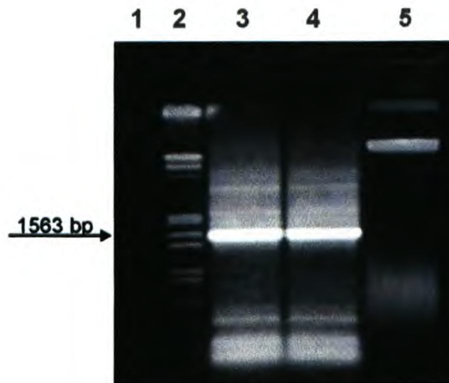

Figure 1.5 Agarose gel electrophoretic analysis of amplified CYP17. Lane 1: Negative control, PCR amplification without DNA template. **Lane 2:** DNA marker λDNA/*Eco*RI/*Hind*III. **Lanes 3 and 4:** Amplified CYP17 (1563 bp) fragment (0.8 μg) using primers STE1 and P104. **Lane 5:** Undigested pFLAG-CYP17 plasmid DNA (0.5 μg) used as template for PCR.

The integrity of the newly constructed pUC18-CYP17 vector was confirmed by restriction digest analyses with *Xba*I followed by 1% agarose gel electrophoresis. The pUC18-CYP17 plasmid (4237bp) with the insert yielded two 3902 and 335 bp fragments as shown in Figure 1.6, lanes 12 and 16. These positive clones were isolated and the presence of the CYP17 was subsequently confirmed by PCR amplification with CYP17 specific primers as shown in Figure 1.7.

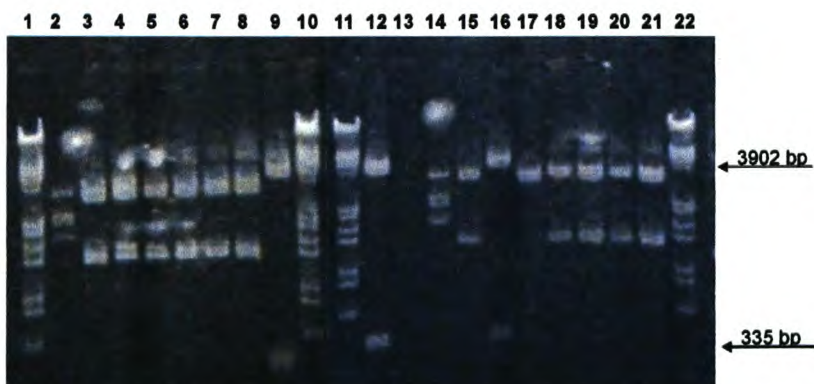

Figure 1.6 Restriction digest analysis of pUC18 plasmids containing the CYP17. Lanes 1, 10, 11 and 22: DNA marker λDNA/*Eco*RI/*Hind*III. **Lanes 9, 12 and 16**: *Xba*I digest of plasmid DNA containing the CYP17 in the 5'- 3' orientation yielding 3902 bp and 335 bp fragments. **Lanes 3-8, 15 and 18-21:** *Xba*I digest of plasmid DNA containing the 3'- 5' orientated CYP17.

Stellenbosch University http://scholar.sun.ac.za

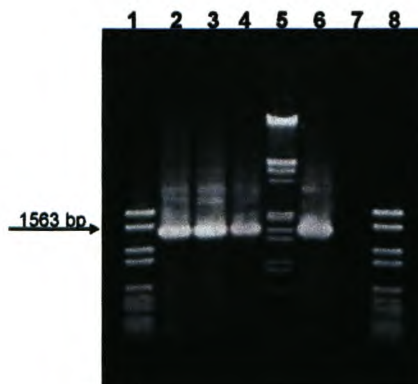

Figure 1.7 PCR amplification analysis of recombinant pUC18-CYP17 plasmids. The reaction products (0.5 µg) were analysed by 1% agarose gel electrophoresis **Lanes 1 and 8**: DNA molecular weight marker VI (Roche). **Lanes 2, 3, 4 and 6:** CYP17 fragment (1563bp) amplified from recombinant vectors using gene specific primers STE1 and P104. **Lane 5:** DNA marker λDNA/*Eco*RI/*Hind*III. **Lane 7:** Negative control, PCR amplification without DNA template.

1.1.2.2 Construction of the pPIC9K-CYP17 expression vector

The *Eco*RI digested pUC18-CYP17 and pPIC9K plasmids were electrophoresed on an 0.8% low melting agarose gel and the CYP17 fragment and linearised plasmid DNA were excised and purified. Prior to ligation the cohesive compatible ends of the *Eco*RI linearised pPIC9K DNA were dephosphorylated and the CYP17 fragment successfully ligated. To evaluate the absence of religated plasmid DNA, the dephosphorylated plasmid (9276 bp) and the CYP17 fragment (1545 bp) were analysed by 1% agarose gel electrophoresis as shown in Figure 1.8.

Figure 1.8 Restriction digest analysis of the CYP17 fragment and pPIC9K plasmid. Lanes 1 and 3: DNA marker λDNA/*Eco*RI/*Hind*III. **Lane 2:** CYP17 fragment (1545 bp) obtained with the *Eco*RI digestion at 0.2 µg. **Lane 4:** The *Eco*RI linearised, dephosphorylated pPIC9K plasmid (1 µg).

Stellenbosch University http://scholar.sun.ac.za

Transformed bacterial colonies, were selected from the LB-ampicillin plates and grown in small-scale cultures. Plasmid DNA isolated from each culture was digested with *Xba*I and analysed by 1% agarose gel electrophoresis. The recombinant pPIC9K-CYP17 vector (10827 bp) with the insert in the correct orientation (5'-3' orientation) should yield 2053 and 8774 bp fragments of as shown in Figure 1.9.

DNA sequencing analyses, Appendix A.1, of these clones confirmed that the CYP17 had been inserted in reading frame for eukaryotic translation initiation and secretion.

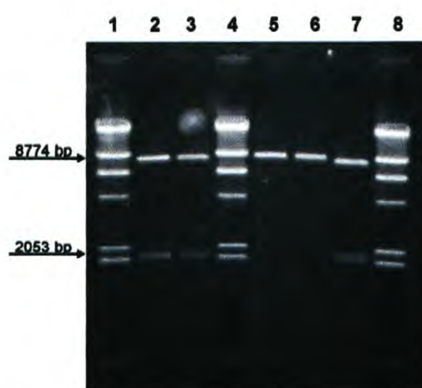

Figure 1.9 Restriction digest analysis of pPIC9K plasmids containing the CYP17. Lanes 1, 4 and 8: DNA marker λDNA/*Hind*III. **Lanes 2, 3 and 7:** *Xba*I digest of pPIC9K vector (8774 bp) containing the CYP17 (2053bp) in the 5'-3' orientation. **Lanes 5 and 6:** *Xba*I digest of pPIC9K vector containing no insert. Each lane contained 0.5 μg DNA.

1.1.2.3 Screening of GS115/pPIC9K-CYP17 and KM71/pPIC9K-CYP17 transformants for integration

The recombinant pPIC9K-CYP17 and the parent vector pPIC9K DNA were linearised with *Sal*I and transformed into two different *P. pastoris* host strains GS115 and KM71. The transformants were selected on the requirement of a His^+ phenotype for growth and selected for single copy integration into the *Pichia* genome by conferring resistancy of G418 at 0.25 mg/ml.

The GS115 His^+ transformants were screened for their methanol utilisation (Mut) phenotype, since methanol will be the sole carbon source used for expression and a wild-type Mut^+ is required. Most of the His^+ transformants should be Mut^+, because transformation of GS115 with *Sal*I-linearised constructs favor recombination at the *HIS4*

Stellenbosch University http://scholar.sun.ac.za

locus. However, with the presence of the *AOX1* gene in the plasmid, there is a chance that recombination will occur at the *AOX1* locus, disrupting the wild-type *AOX1* gene and creating His^+Mut^s transformants. The deletion of the *AOXI* results in the loss of the 2100 bp PCR amplified fragment.

PCR amplification followed by 1% agarose gel electrophoretic analyses of lysed His^+ transformants could not determine if the CYP17 was integrated correctly, since the expected two fragments have similar sizes and could not be separated. One corresponded to the CYP17 (2044 bp) and the other to the *AOX1* gene (2100 bp) as shown in Figure 1.10, lane 3. A gene-specific primer, *STE1* together with 5' AOX1P primer was used to confirm the presence of CYP17 (1680 bp) as shown Figure 1.10, lane 4. Mut^+ integrants with only the parent vector pPIC9K, produced 492 bp and 2100 bp PCR products as seen in Figure 1.10, lane 5.

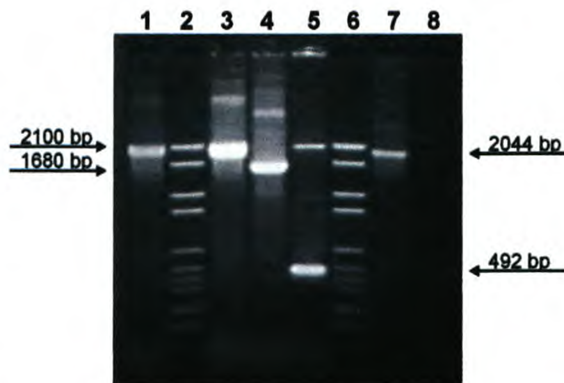

Figure 1.10 Direct PCR screening of GS115/pPIC9K-CYP17 clones. Lane 1: Amplification of the wild-type *AOX1* gene (2100 bp) using AOX1P primers. **Lanes 2 and 6:** DNA molecular weight marker VI. **Lanes 3 and 4:** Contains a clone carrying both the *AOX1* gene (2100 bp) and CYP17 (2044 bp). In lane 3, AOX1P primers were used and in lane 4 an insert-specific primer STE1 and the 3'AOX1P primer were used resulting in a 1680 bp PCR product. **Lane 5:** Contains a clone with the *AOX1* gene(2100 bp) carrying the pPIC9K vector without insert (492 bp). AOX1P primers were used. **Lane 7:** Positive PCR control containing CYP17 fragment (2044 bp) amplified from pPIC9K-CYP17 plasmid using AOX1P primers. **Lane 8:** Negative control, PCR amplification without DNA template. Each lane contained 20 µl reaction product.

The successful integration of the CYP17 and truncCYP17 into the yeast genomic DNA was shown by *Sac*I restriction digestion of the amplified PCR products. The integrated CYP17, with two *Sac*I digestion sites, produced four fragments. Three corresponded to the CYP17 (1344, 385 and 315 bp) and the remaining to the *AOX1* gene (2100 bp) as shown in Figure 1.11, lane 5. The truncCYP17 has only one *Sac*I digestion site since the

Stellenbosch University http://scholar.sun.ac.za

second site was in the 5'-terminal of the full-length CYP17, which had been excised. Therefore three bands were produced. Two corresponded to the truncCYP17 (1675 and 315 bp) and the other to the *AOX1* gene (2100 bp). This was confirmed by 1% agarose gel electrophoresis (Figure 1.11, lane 7).

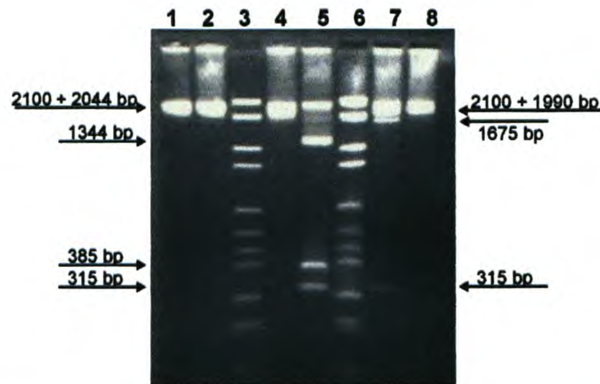

Figure 1.11 Restriction digest analysis of the PCR products from GS115/pPIC9K-CYP17 and GS115/pPIC9K-truncCYP17 clones. Lanes 1 and 2: undigested and *Sac*I digested PCR products (0.5 μg) of the wild-type *AOX1* gene, respectively. **Lanes 3 and 6:** DNA molecular weight marker VI. **Lanes 4 and 5:** undigested and *Sac*I digested PCR products (0.5 μg) of the clone carrying both the *AOX1* gene and the CYP17, respectively. *Sac*I digest resulted in four bands. The 1344, 385 and 315 bp fragments corresponded to the CYP17 and the 2100 bp band to the *AOX1* gene. **Lanes 7 and 8:** *Sac*I digested and undigested PCR products (0.5 μg) of the clone carrying both the *AOX1* gene and the truncCYP17, respectively. *Sac*I digest resulted in three bands. The 1675 and 315 bp fragments corresponded to the truncCYP17 and the 2100 bp band to the *AOX1* gene.

The same PCR analysis was carried out on KM71 integrants. The results, analysed by 1% agarose gel electrophoresis, were similar except that the wild-type *AOX1* gene in KM71 cells was approximately 4500 bp instead of 2100 bp, because the *ARG4* was inserted into the *AOX1* gene. Two bands were produced. One corresponded to the CYP17 (2044 bp) and the other to the *AOX1* gene (4500 bp) as shown in Figure 1.12, lane 1.

Stellenbosch University http://scholar.sun.ac.za

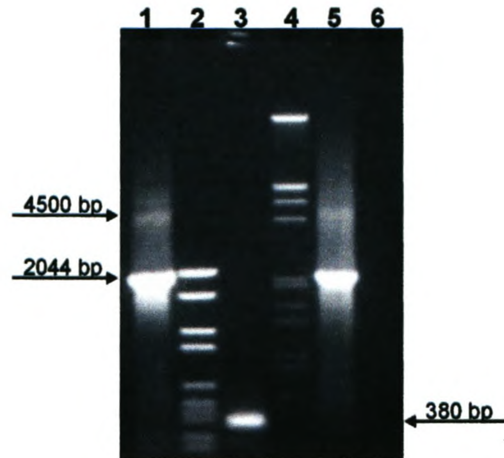

Figure 1.12 Direct PCR screening of KM71/pPIC9K-CYP17 clones. The reaction products (20 µl) were analysed by 1% agarose gel electrophoresis. Lanes 1 and 3: Contains a clone carrying both the wild-type *AOX1* gene (4500 bp) and the CYP17 insert cloned into the pPIC9K plasmid resulting in a 2044 bp PCR product. In lane 1 AOX1P primers were used and in lane 3 two insert-specific primers STE1 and 3989 were used resulting in a 380 bp PCR product. Lane 2: DNA molecular weight marker VI. Lane 4: DNA marker λDNA/*Eco*RI/*Hind*III. Lane 5: Positive PCR control containing CYP17 fragment (2044 bp) amplified from pPIC9K-CYP17 plasmid using AOX1P primers. Lane 6: Negative control, PCR amplification without DNA template.

Stellenbosch University http://scholar.sun.ac.za

1.2 Cloning and construction of the truncated human CYP17 (truncCYP17) into the pPIC9K expression vector

1.2.1 Materials and methods

1.2.1.1 Subcloning of the truncCYP17 cDNA

The full-length human CYP17 in the pPIC9K-CYP17 construct was used as the template DNA to amplify a truncCYP17 cDNA, depicted in Figure 1.13. An upstream primer P103 (5'-ATGCGAATTCAAGAGAAGGTGCCCTGGTGC-3') was designed to delete the first 18 amino acid residues (membrane binding site) and the downstream primer P104 complementary to the histidine tag including the stop codon. Both primers introduced *Eco*RI sites (underlined).

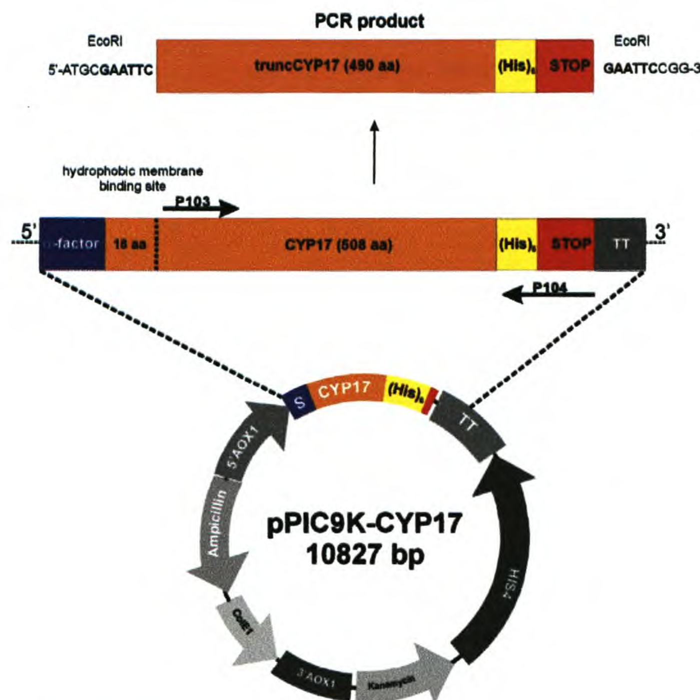

Figure 1.13 Schematic representation of the cloning strategy for the truncCYP17 amplified from the pPIC9K-CYP17 expression vector. The nucleotides encoding the first 18 amino acid residues from the 5'-terminal of the CYP17 were truncated using the upstream primer P103, which is complementary to the DNA sequence downstream from the nucleotides encoding the first 18 amino acid residues. The downstream primer P104 is complementary to the 3'-terminal including the stop codon and nucleotide bases encoding six histidine residues. Both primers include a nucleotide sequence encoding an *Eco*RI site for subcloning into the pUC18 vector.

Stellenbosch University http://scholar.sun.ac.za

Each PCR amplification mixture (100 µl) contained pPIC9K-CYP17 vector DNA (750 ng) and the reaction was carried out as described in section 1.1.1.2. The amplified DNA fragment was electrophoresed on an 0.8% low melting agarose gel, extracted and purified as described in section 1.1.1.2.

The truncCYP17 PCR product was digested with *Eco*RI resulting in a fragment containing *Eco*RI digestion sites on the 5' and 3'-termini which subsequently was cloned into the pUC18 vector linearised with *Eco*RI as shown in Figure 1.14. Two separate restriction digestions, one containing amplified truncCYP17 DNA (2 µg) and the other containing pUC18 DNA (2 µg) were carried out as described in section 1.1.1.2.2. The digested CYP17 fragment and linearised vector were electrophoresed on an 0.8% low-melting agarose gel in TAE buffer (40 mM) and purified as described in section 1.1.1.2.1.

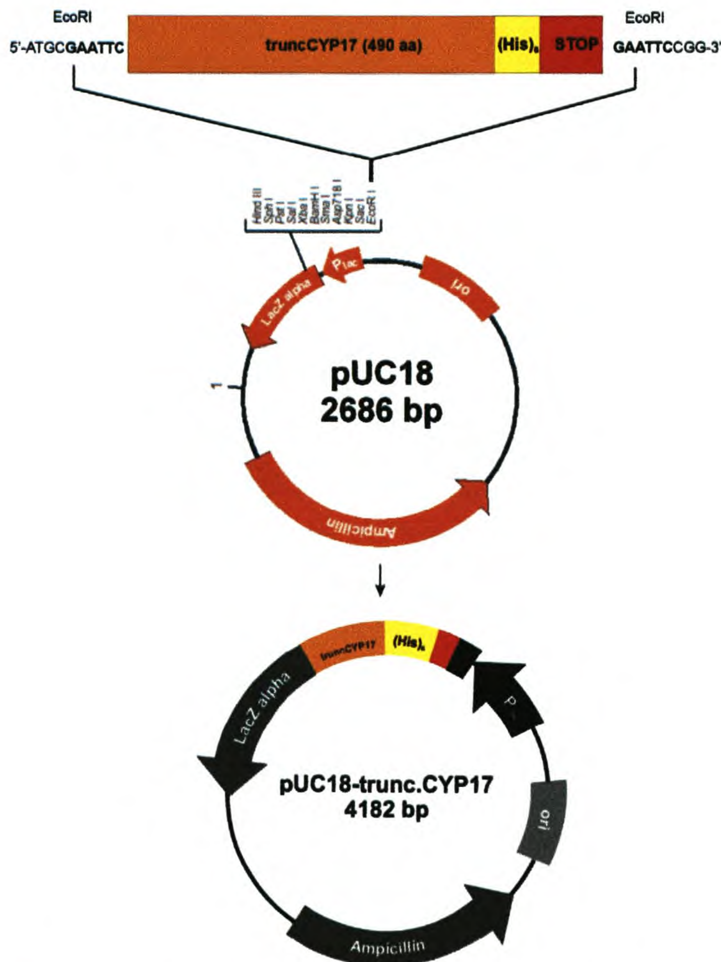

Figure 1.14 Schematic representation of the subcloning of the truncCYP17 into pUC18 vector.

Stellenbosch University http://scholar.sun.ac.za

The linearised pUC18 vector (1.89 μg) was dephosphorylated with CIAP (0.022 U) as described in section 1.1.1.2.3. The purified digested truncCYP17 fragment (112 ng) and pUC18 DNA (200 ng) were subsequently ligated as described in section 1.1.1.2.4. Competent JM109 cells were transformed with the ligated vectors. In addition to ampicillin selection, the vector contained the lacZ gene, which allowed for blue/white colony selection.

Positive colonies were subsequently incubated overnight in LB medium (5 ml) at 37°C and the plasmid DNA was isolated using the Wizard™ *Plus* Minipreps DNA Purification System as described in section 1.1.1.2.5. The purified recombinant plasmid DNA was digested with *Xba*I and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit as described in section 1.1.1.2.5. The purified isolated plasmids were analysed by DNA sequencing and subsequently used in the construction of the pPIC9K-truncCYP17 expression vector.

1.2.1.2 Construction of the pPIC9K-truncCYP17 expression vector

The pUC18-truncCYP17 construct (4182 bp), was digested with *Eco*RI and the truncCYP17 was ligated into pPIC9K DNA linearised with *Eco*RI as shown in Figure 1.15. Two separate restriction digestions, one containing pUC18-CYP17 (2 μg) and the other pPIC9K DNA (2 μg) were carried out at 37°C for 90 min in a final volume of 30 μl containing 5.0 μl restriction buffer B (supplied), acetylated BSA (5.0 μg) and *Eco*RI (10 U). The digested truncCYP17 fragment and the linearised vector were electrophoresed on an 0.8% low-melting agarose gel and extracted from the agarose gel as described in section 1.1.1.2.1.

The linearised pPIC9K DNA (14μg) was dephosphorylated with CIAP (0.05 U) as described in section 1.1.1.2.3. The purified digested truncCYP17 DNA (32.3 ng) and pPIC9K plasmid DNA (200 ng) were subsequently ligated as described in section 1.1.1.2.4. Competent JM109 cells were transformed with the ligated vectors and transformants were selected on LB-agar plated with ampicillin as described in section 1.1.1.4.1.

Stellenbosch University http://scholar.sun.ac.za

Positive colonies were incubated overnight in LB medium (5 ml) at 37°C and the plasmid DNA was isolated as described in section 1.1.1.2.5. The purified recombinant plasmids were digested with *Xba*I and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit as described in section 1.1.1.2.5.

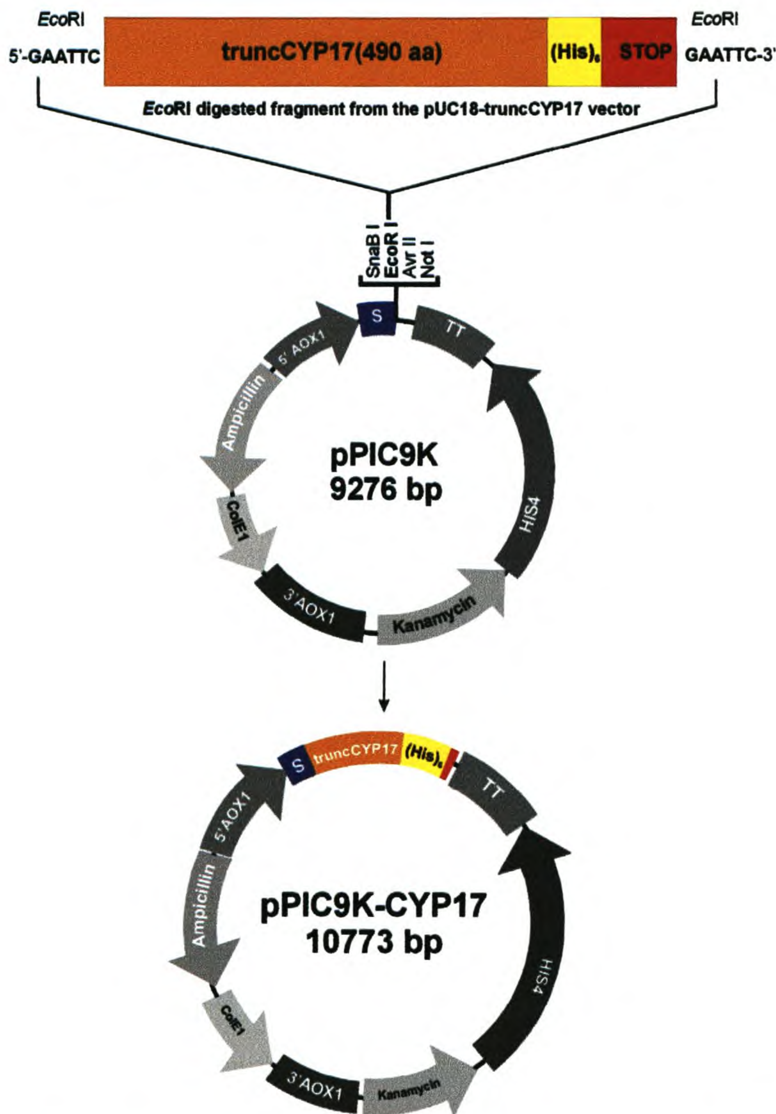

Figure 1.15 Schematic representation of the cloning strategy for the construction of the pPIC9K-truncCYP17 expression vector. The cloned truncCYP17 in the pUC18 vector was digested with *Eco*RI and inserted into the pPIC9K expression vector at the *Eco*RI site. The truncCYP17 was cloned in reading frame of the α-factor secretion signal and contains a nucleotide sequence encoding the six histidine tag.

Stellenbosch University http://scholar.sun.ac.za

The recombinant pPIC9K-truncCYP17 vector was analysed by DNA sequencing as described in section 1.1.1.4.3. The truncated CYP17 with its insertion sites was sequenced with seven primers as illustrated in Figure 1.16. The DNA sequencing analyses was carried out as described in 1.1.1.4.3, confirming correct insertion.

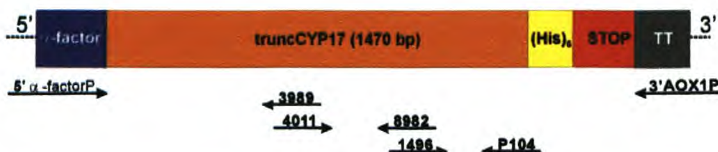

Figure 1.16 Schematic representation of the DNA sequencing strategy for the truncCYP17 in the pPIC9K-truncCYP17 recombinant vector.

1.2.1.3 Screening of GS115/pPIC9K-truncCYP17 and KM71/pPIC9K-truncCYP17 transformants for integration

The cloned expression vectors, pPIC9K-truncCYP17 (10 μg) and the parent vector pPIC9K (10 μg), were linearised with *Sal*I and transformed into competent *P. pastoris* strains, GS115 Mut$^+$ (Methanol utilisation plus) and KM71 Mut^s (Methanol utilisation slow) using the spheroplast method as described in section 1.1.1.5. The presence of truncCYP17 cDNA integrated into the yeast chromosomes was subsequently confirmed by PCR amplification analyses as described in section 1.1.1.5.

1.2.2 Results

1.2.2.1 Subcloning of the truncCYP17 cDNA

The full-length human CYP17, located in the pPIC9K-CYP17 expression vector was used as the template to amplify a truncCYP17 cDNA, yielding a 1510 bp product. This amplified product was subsequently subcloned into an intermediate vector, pUC18, before cloning it into the pPIC9K expression vector.

The integrity of the newly constructed pUC18-truncCYP17 vector was confirmed by restriction digest analyses with *Xba*I followed by 1% agarose gel electrophoresis. The pUC18-truncCYP17 plasmid (4182 bp), containing the insert yielded two 3902 and 280 bp fragments as shown in Figure 1.17, lanes 7, 10 and 11.

Stellenbosch University http://scholar.sun.ac.za

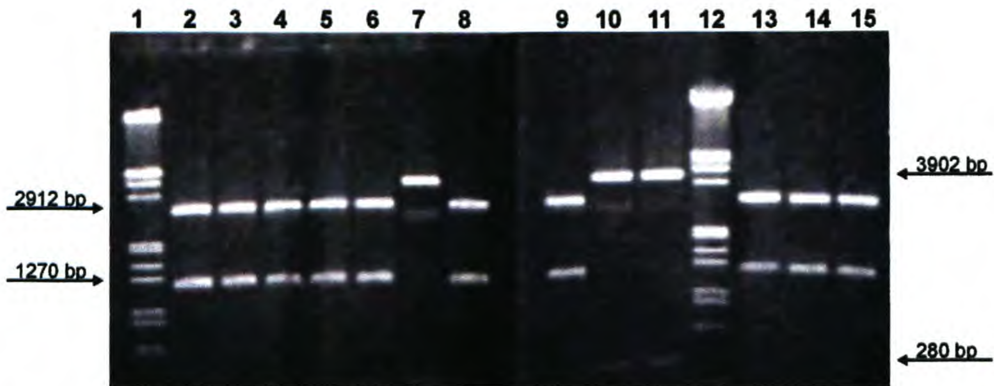

Figure 1.17 Restriction digest analysis of pUC18 plasmids containing the truncCYP17. Lanes 1 and 12: DNA marker λDNA/*Eco*RI/*Hind*III. **Lanes 2-6, 8, 9, and 13-15**: *Xba*I digest of plasmid DNA (0.5 μg) containing the truncCYP17 in the 3'-5' orientation yielding 2912 and 1270 bp fragments. **Lanes 7, 10 and 11:** *Xba*I digest of pUC18 plasmid DNA (0.5 μg) containing the truncCYP17 in the 5'-3' orientation yielding 3902 and 280 bp fragments.

1.2.2.2 Construction of the pPIC9K-truncCYP17 expression vector

The *Eco*RI digested pUC18-truncCYP17 and pPIC9K DNA were electrophoresed on an 0.8% low melting agarose gel and the truncCYP17 and linearised plasmid DNA were excised and purified. Prior to ligation the cohesive compatible ends of the linearised pPIC9K DNA were dephosphorylated and the truncCYP17 successfully ligated.

Transformed bacterial colonies, were grown in small-scale cultures. Plasmid DNA isolated from each culture was digested with *Xba*I and analysed by 1% agarose gel electrophoresis. The cloned pPIC9K-truncCYP17 vector (10773 bp) with the insert in the correct orientation (5'-3' orientation) should yield 2053 and 8720 bp fragments as shown in Figure 1.18, lanes 13 and 16.

Analyses of PCR amplifications, seen in Figure 1.19 and DNA sequencing analyses, Appendix A.2, confirmed that the truncCYP17 had been inserted in reading frame for eukaryotic translation initiation and secretion.

1.2.2.3 Screening of GS115/pPIC9K-truncCYP17 transformants for integration

Results were shown in Figure 1.11.

Stellenbosch University http://scholar.sun.ac.za

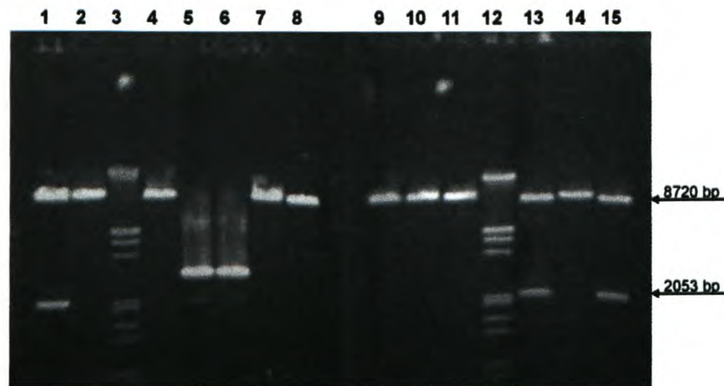

Figure 1.18 Restriction digest analysis of pPIC9K plasmids containing the truncCYP17. Lane 1: Positive control containing *Xba*I digest of pPIC9K-CYP17 plasmid (0.5 μg). **Lanes 2,4-11 and 14:** *Xba*I digest of plasmid DNA (0.5 μg) without the truncCYP17. **Lanes 3 and 12:** DNA marker λDNA/*Eco*RI/*Hind*III. **Lanes 13 and 15:** *Xba*I digest of plasmid DNA (0.5 μg) containing the truncCYP17 in the 5'-3' orientation yielding 8720 and 2053 bp fragments.

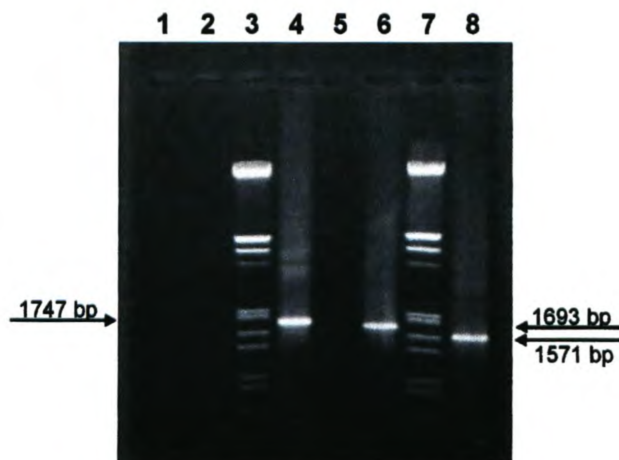

Figure 1.19 PCR analysis of the recombinant pPIC9K-CYP17 and pPIC9K-truncCYP17 plasmids. The reaction products (20 μl) were analysed by 1% agarose gel electrophoresis. **Lanes 1 and 2:** Negative controls, PCR amplifications without DNA template. **Lanes 3 and 7:** DNA marker λDNA/*Eco*RI/*Hind*III. **Lane 4:** Amplified CYP17 (1747 bp). **Lane 6:** Amplified truncCYP17 (1693 bp). All amplifications were carried out with 5'-α-factorP and 3'-AOX1P primers. **Lane 8:** Positive control containing amplified pPIC9K-CYP17 plasmid using primers P105 and P106.

Stellenbosch University http://scholar.sun.ac.za

1.3 Cloning and construction of the full-length human b_5 gene (CYB5) into the pPIC9K expression vector

1.3.1 Materials and Methods

1.3.1.1 Subcloning of the CYB5 cDNA

All DNA manipulations were carried out using standard procedures (55). The full-length human CYB5, located in the pCMV5-CYB5 vector[1], was used as the DNA template to amplify a CYB5 cDNA as depicted in Figure 1.20. The first amplification used the upstream primer P107 (5'-AGGCGAATTCATGGCAGAGCAGTCGGAC-3') and the downstream primer P108 (5'-**GTGATGGTGATGGTGATG**GTCCTCTGCCATG TATAGG-3'). The former introduced an *Eco*RI site (underlined) and the latter a six histidine coding region (bold). The PCR 1 product was the template for the second amplification using the forward primer P109 (5'-AGGCGAATTCATGGCAGA-3') and reverse primer P110 (5'-CCGGAATTC*TCA***GTGATGGTGATGGTGATG**GT-3'). The latter introduced a stop codon (italics) after the nucleotide bases encoding six histidine residues followed by an *Eco*RI site.

Each PCR amplification mixture (100 μl) contained pCMV5-CYB5 vector DNA (750 ng) and the reaction was carried out as described in section 1.1.1.2. The PCR 1 product was electrophoresed on an 0.8% low melting agarose gel, extracted and purified as described in section 1.1.1.2. The same reaction procedure was followed for a second amplification using the purified PCR 1 product (400 ng) as a template. The PCR 2 product was electrophoresed, extracted and purified as described in section 1.1.1.2.

The CYB5 PCR 2 product was digested with *Eco*RI and cloned into the pUC18 vector linearised with *Eco*RI, depicted in Figure 1.21. Two separate restriction digestions, one containing amplified CYB5 DNA (2 μg) and the other containing pUC18 DNA (2 μg) were carried out as described in section 1.1.1.2.2.

---

[1] A kind gift from Prof Ronald Estabrook, University of Texas Medical Centre, Dallas, Texas, USA

Stellenbosch University http://scholar.sun.ac.za

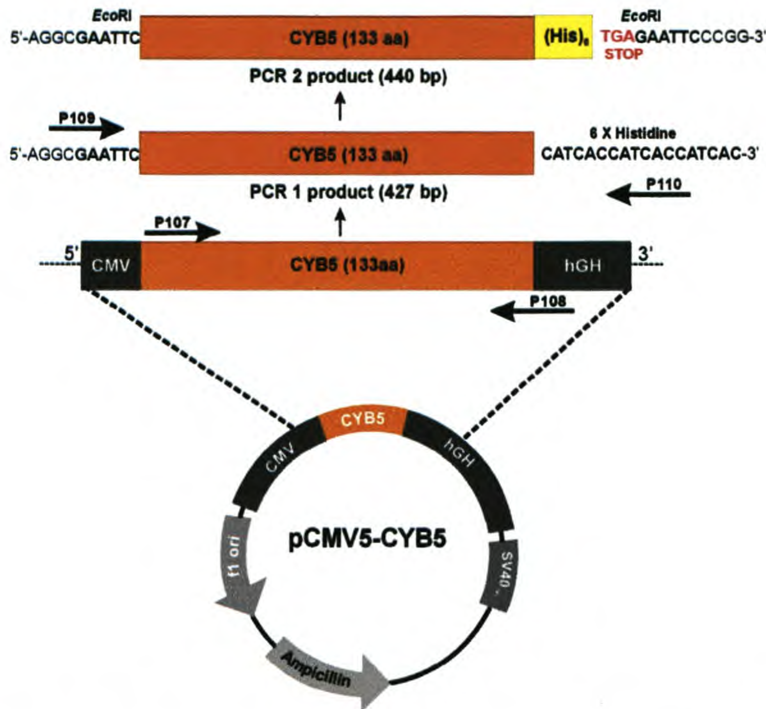

Figure 1.20 Schematic representation of the cloning strategy for the CYB5 amplified from the pCMV5-CYB5 vector. The first PCR amplification using the upstream primer P107 and the downstream primer P108 introduced a *Eco*RI site and a nucleotide sequence encoding six histidine residues, respectively. P109 and P110 primers were used in the second amplification. The latter introduced nucleotide bases encoding a stop site followed by an *Eco*RI site.

The digested CYB5 fragment and linearised vector were electrophoresed on an 0.8% low-melting agarose gel in 40 mM TAE buffer and purified as described in section 1.1.1.2.1. The linearised pUC18 vector (3.66 μg) was dephosphorylated with CIAP (0.042 U) as described in section 1.1.1.2.3. The purified digested CYB5 fragment (62.4 ng) and pUC18 (200 ng) DNA were subsequently ligated as described in section 1.1.1.2.4. Competent JM109 cells were transformed with the ligated vectors. In addition to ampicillin selection, the vector contained the lacZ gene which allowed for blue/white colony selection.

Positive colonies were subsequently incubated overnight in LB medium (5 ml) at 37°C and the plasmid DNA was isolated using the Wizard™ *Plus* Minipreps DNA Purification System as described in section 1.1.1.2.5. The purified recombinant plasmid DNA was digested with *Pst*I / *Eco*RI and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit as described in section

Stellenbosch University http://scholar.sun.ac.za

1.1.1.2.5. The purified isolated plasmids were analysed by DNA sequencing and subsequently used in the construction of the pPIC9K-CYB5 expression vector.

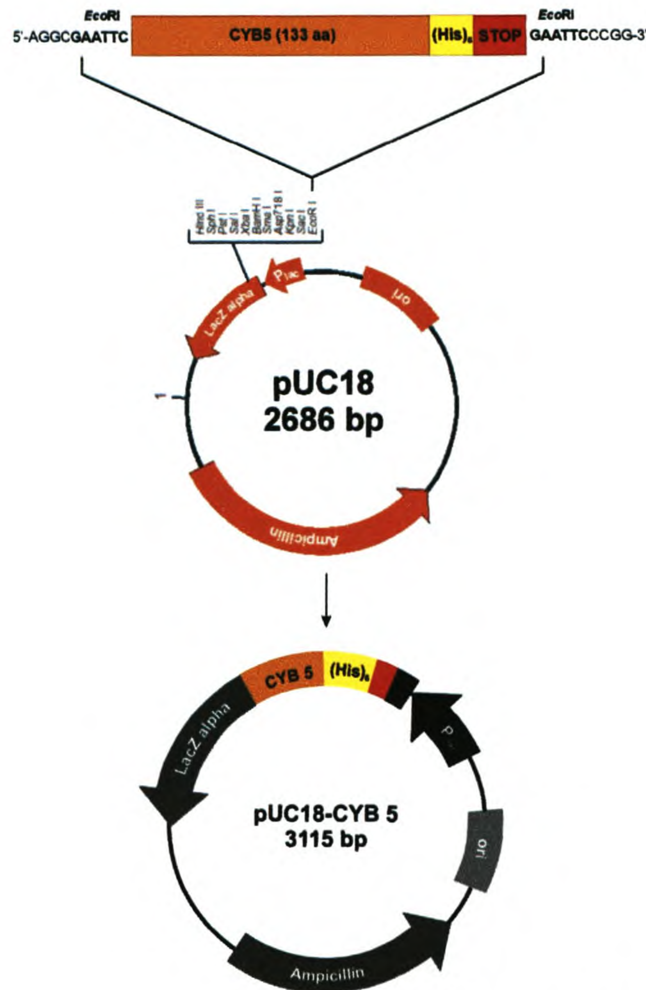

Figure 1.21 Schematic representation of the subcloning of the CYB5 into pUC18 vector.

1.3.1.2 Construction of the pPIC9K-CYB5 expression vector

The pUC18-CYB5 construct (3115 bp), was digested with *Eco*RI and the CYB5 was ligated into pPIC9K vector, linearised with *Eco*RI as depicted in Figure 1.22. Two separate restriction digestions, one containing pUC18-CYB5 (2 μg) and the other pPIC9K DNA (2 μg) were carried at 37°C for 2 hours, each in a final volume of 30 μl containing 5.0 μl restriction buffer B (supplied), acetylated BSA (5.0 μg) and *Eco*RI (10 U). The digested CYB5 fragment and linearised vector were electrophoresed on an 0.8% low-melting agarose gel, extracted and purified from the agarose as described in section 1.1.1.2.1.

Stellenbosch University http://scholar.sun.ac.za

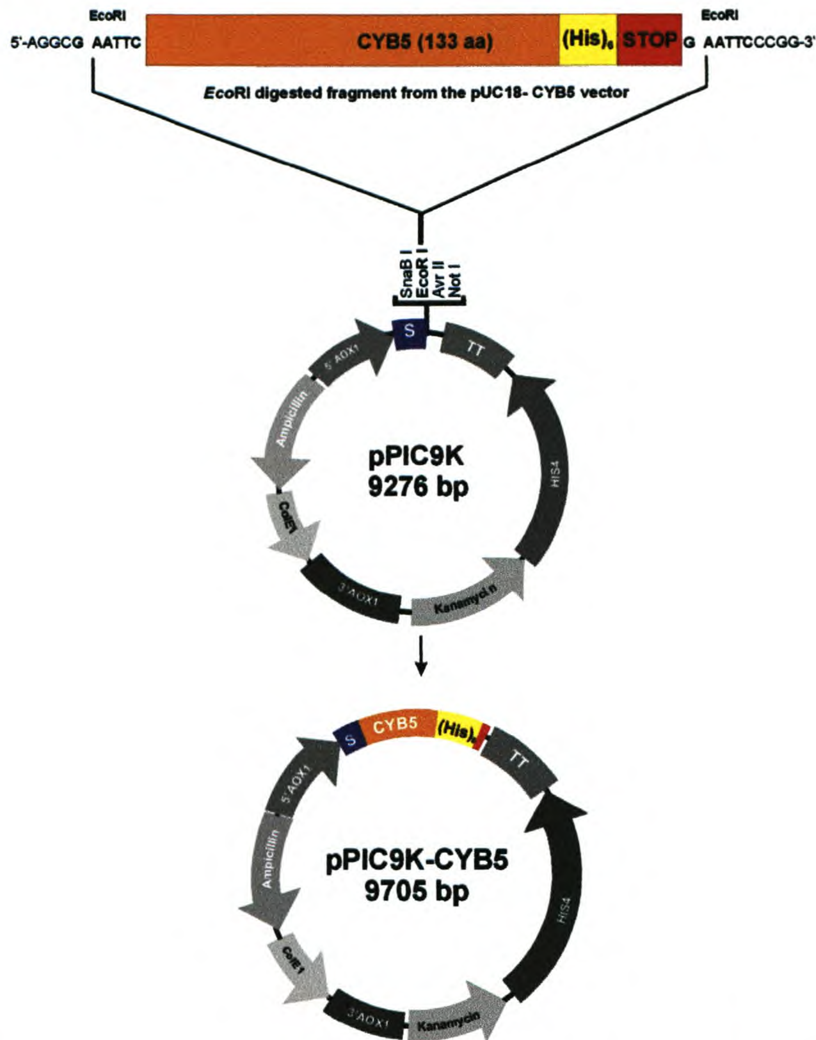

Figure 1.22 Schematic representation of the cloning strategy for the construction of the pPIC9K-CYB5 expression vector. The cloned CYB5 in the pUC18 vector was digested with *Eco*RI and inserted into the pPIC9K expression vector at the *Eco*RI *site*. The CYB5 was cloned in reading frame of the α-factor secretion signal and contains a nucleotide sequence encoding the six histidine tag.

The linearised pPIC9K DNA (4.62 μg) was dephosphorylated with CIAP (0.015 U) and purified as described in section 1.1.1.2.3. The purified digested CYB5 (8.71 ng) and pPIC9K DNA (200 ng) were subsequently ligated as described in section 1.1.1.2.4. Competent JM109 cells were transformed with the ligated vectors and transformants were selected on LB-agar plates with ampicillin as described in section 1.1.1.4. Positive colonies were subsequently incubated overnight in LB medium (5 ml) at 37°C and the plasmid DNA was isolated as described in section 1.1.1.2.5. The purified recombinant plasmid DNA was digested with *Sac*I and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified

Stellenbosch University http://scholar.sun.ac.za

alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit as described in section 1.1.1.2.5.

The purified recombinant pPIC9K-CYB5 vector was analysed by restriction digestion and DNA sequence analyses as described in 1.1.1.4.3. Primers 5'AOX1P and 3'AOX1P were used as illustrated in Figure 1.23.

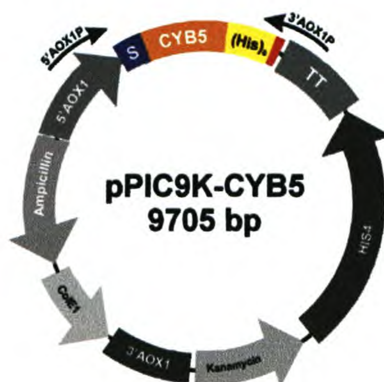

Figure 1.23 Schematic representation of the DNA sequencing strategy for the CYB5 in the pPIC9K-CYB5 expression vector.

1.3.1.3 Screening of GS115/pPIC9K-CYB5 and KM71/pPIC9K-CYB5 transformants for integration

The cloned expression vectors, pPIC9K-CYB5 (10 μg) and the parent vector pPIC9K (10 μg), were linearised with *Sal*I and transformed into competent *P. pastoris* strains, GS115 Mut^+ (Methanol utilisation plus) and KM71 Mut^s (Methanol utilisation slow) using the spheroplast method as described in section 1.1.1.5. The presence of CYB5 cDNA integrated into the yeast chromosomes was subsequently confirmed by PCR amplification analyses as described in section 1.1.1.5.

1.3.2 Results

1.3.2.1 Subcloning of the CYB5 cDNA

The full-length human CYB5, located in the pCMV5-CYB5 expression vector was used as the template to amplify a CYB5 cDNA, yielded a PCR 1 product (427 bp). This amplified product was subsequently used as a template for the second amplification,

Stellenbosch University http://scholar.sun.ac.za

which yielded the expected PCR 2 product (440 bp) as seen in Figure 1.24. PCR 2 product was subcloned into an intermediate vector, pUC18, before cloning it into the pPIC9K expression vector.

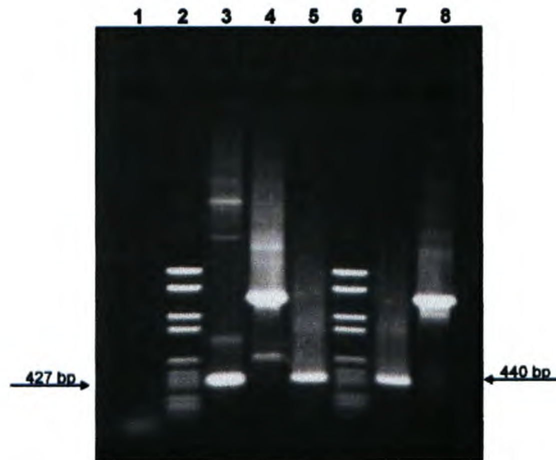

Figure 1.24 Agarose gel electrophoretic analysis of amplified CYB5. The reaction products (10 µl) were analysed by 1% agarose gel electrophoresis. **Lane 1**: Negative control, PCR amplification without DNA template. **Lanes 2 and 6:** DNA molecular weight marker VI. **Lane 3:** Amplified CYB5 (PCR 1 product, 427 bp using primers P107 and P108. **Lanes 5 and 7:** Amplified CYB5 (PCR 2 product, 440 bp) using primers P109 and P110. **Lanes 4 and 8:** Positive PCR controls containing CYP17 fragment amplified from pPIC9K-CYP17 using primers P105 and P106.

The integrity of the newly constructed pUC18-CYB5 vector was confirmed by restriction digest analyses with either *Pst*I or *Eco*RI followed by 1% agarose gel electrophoresis. The pUC18-CYB5 plasmid (3115 bp) containing the insert yielded the expected two bands corresponding to 2715 and 400 bp for the *Pst*I digestion and two bands corresponding to 2689 and 426 bp for the *Eco*RI digestion as shown in Figure 1.25, lanes 3 and 11.

1.3.2.2 Construction of the pPIC9K-CYB5 expression vector

The *Eco*RI digested pUC18-CYB5 and pPIC9K plasmid DNA were electrophoresed on an 0.8% low melting agarose gel and the CYB5 and linearised plasmid were excised and purified. Prior to ligation the cohesive compatible ends of the linearised pPIC9K plasmid DNA were dephosphorylated and the CYB5 successfully ligated.

Stellenbosch University http://scholar.sun.ac.za

Transformed bacterial colonies were grown in small-scale cultures. Plasmid DNA isolated from each culture was digested with *Sac*I and analysed by 1% agarose gel electrophoresis. The cloned pPIC9K-CYB5 vector (9705 bp) yielded the correct

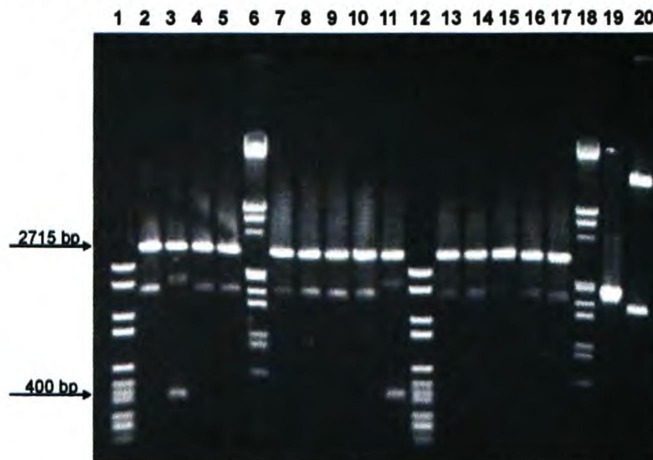

Figure 1.25 Restriction digest analysis of pUC18 plasmids containing the CYB5. Lanes 1 and 12: DNA molecular weight marker VI. **Lanes 2, 4, 5, 7, 8 and 9:** *Pst*I digest of plasmid DNA (0.5 μg) without insert. **Lane 3:** *Pst*I digest of plasmid DNA containing the CYB5 yielding 2715 and 400 bp fragments. **Lanes 6 and 18:** DNA marker λDNA/*Eco*RI/*Hind*III. **Lanes 10, 13-17:** *Eco*RI digest of plasmid DNA (0.5 μg) without insert. **Lane 11:** *Eco*RI digest of plasmid DNA (0.5 μg) containing the CYB5. **Lane 19:** undigested plasmid DNA (0.5 μg). **Lane 20:** Positive control containing *Eco*RI digested pPIC9K-CYP17 plasmid DNA (0.5 μg).

8435 and 1270 bp fragments as shown in Figure 1.26, lanes 3, 14 and 19.

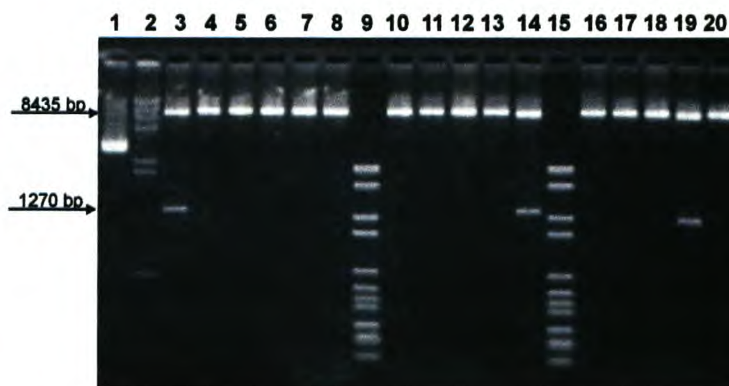

Figure 1.26 Restriction digest analysis of pPIC9K plasmids containing the CYB5. Lane 1: Positive control containing *Sac*I digested pPIC9K plasmid DNA (0.5 μg). **Lane 2:** DNA marker λDNA/*Hind*III. **Lanes 9 and 15:** DNA molecular weight marker VI. **Lanes 3, 14 and 19:** *Sac*I digest of plasmid DNA (0.5 μg) containing the CYB5 in the 5'-3' orientation yielding 8435 and 1270 bp fragments. **Lanes 4-8, 10-13, 16-18 and 20:** *Sac*I digest of plasmid DNA (0.5 μg) without an insert.

The 1% agarose gel electrophoresis could not distinguish between the correctly 5'-3' orientated (8435 and 1270 bp) and incorrectly 3'-5' orientated (8518 and 1187 bp) insert

Stellenbosch University http://scholar.sun.ac.za

in the plasmid. Therefore a 2% agarose gel electrophoresis, separating and resolving the smaller sized DNA, was used to identify the correct fragments. Out of the three digested plasmids, shown in Figure 1.26, lanes 3, 14 and 19, the first two clones were in the correct orientation as shown in Figure 1.27, lanes 2 and 4. DNA sequence analyses, Appendix A.3, of these clones confirmed that the CYB5 had been inserted in reading frame for eukaryotic translation initiation and secretion.

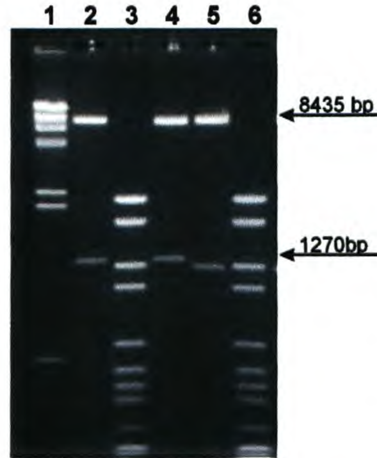

Figure 1.27 Restriction digest analysis of pPIC9K plasmids containing CYB5. The reaction products (0.5 μg) were analysed by 1% agarose gel electrophoresis. **Lane 1:** DNA marker λDNA/*Hind*III. **Lanes 2 and 4:** *Sac*I digest of plasmid DNA containing the CYB5 in the 5'-3' orientation yielding 8435 and 1270 bp fragments. **Lanes 3 and 6:** DNA molecular weight marker VI. **Lane 5:** *Sac*I digest of plasmid DNA containing the CYB5 in the incorrect (3'-5') orientation.

1.3.2.3 Screening of GS115/pPIC9K-CYB5 and KM71/pPIC9K-CYB5 transformants for integration

PCR followed by 1% agarose gel electrophoretic analyses of lysed His^+ transformants showed that the CYB5 was integrated correctly, since two bands were produced. One corresponded to the CYB5 (429 bp) and the other to *AOX1* gene (2100 bp) as shown in Figure 1.28, lane 3. A gene-specific primer P107 together with 5'-AOX1P primer (Figure 1.28, lane 4) was also used to confirm the presence of CYB5. Mut^+ integrants with only the parent vector pPIC9K produced a 492 and 2100 bp PCR products as seen in Figure 1.28, lane 5.

Stellenbosch University http://scholar.sun.ac.za

The same PCR analysis was carried out on KM71 integrants. Two bands were produced. One corresponded to the CYB5 (921 bp) and the other to the *AOX1* gene (4500 bp) as shown in Figure 1.29, lane 3.

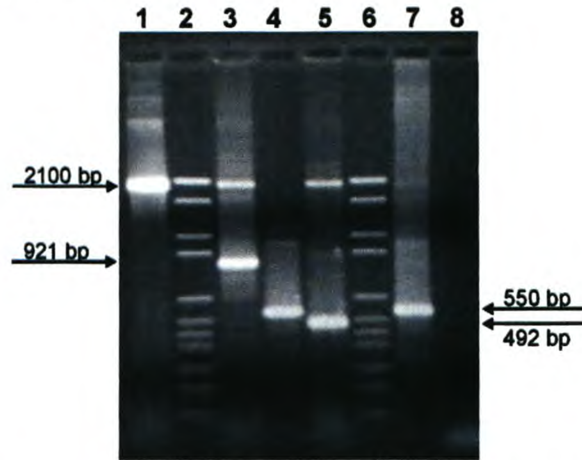

Figure 1.28 Direct PCR screening of GS115/pIC9K-CYB5 clones. Lane 1: Amplification of the wild-type *AOX1* gene (2100 bp) using AOX1P primers. **Lanes 2 and 6:** DNA molecular weight marker VI. **Lanes 3 and 4:** Contains a clone carrying both the *AOX1* gene (2100 bp) and CYB5 insert cloned into the pPIC9K plasmid resulting in a 921 bp PCR product. In lane 3'AOX1P primers were used and in lane 4 insert-specific primer P107 and the 3'AOX1P primer were used resulting in a 550 bp PCR product. **Lane 5:** Contains a clone with the *AOX1* gene (2100 bp) carrying the pPIC9K vector without insert (492 bp). AOX1P primers were used. **Lane 7:** Positive PCR control containing CYPB fragment (550 bp) amplified from pPIC9K-CYB5 plasmid using AOX1P primers. **Lane 8:** Negative control, PCR amplification without DNA template. 20 μl reaction mixtures were loaded.

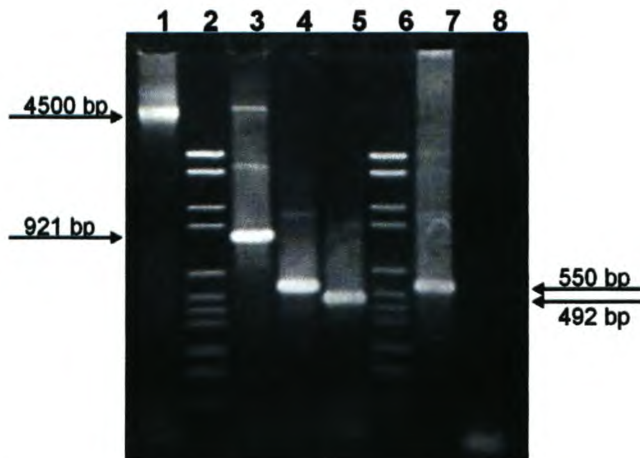

Figure 1.29 Direct PCR screening of KM71/pIC9K-CYB5 clones. The reaction product (20 μl) were analysed by 1% agarose gel electrophoresis. **Lane 1:** Amplification of the wild-type *AOX1* gene (4500 bp) using AOX1P primers. **Lanes 2 and 6:** DNA molecular weight marker VI. **Lanes 3 and 4:** Contains a clone carrying both the *AOX1* gene (4500 bp) and the CYB5 (921 bp). In lane 3 AOX1P primers were used and in lane 4 a gene specific primer P107 and the 3'AOX1P primer were used resulting in a 550 bp fragment. **Lane 5:** Contains a clone carrying the parent pPIC9K vector without insert (492 bp). AOX1P primers were used. **Lane 7:** Positive PCR control containing CYB5 fragment (550 bp) amplified from pPIC9K-CYB5 plasmid using AOX1P primers. **Lane 8:** Negative control, PCR amplification without DNA template. 20 μl reaction mixtures were loaded.

Stellenbosch University http://scholar.sun.ac.za

1.4 Cloning and construction of full-length human CYP19 into the pPIC9K expression vector

1.4.1 Materials and Methods

1.4.1.1 Subcloning of the CYP19 cDNA

All DNA manipulations were carried out using standard procedures (55). The full-length human CYP19, located in the pCMV-CYP19 vector[2], was used as the DNA template to amplify a CYP19 cDNA as depicted in Figure 1.30. The first amplification used the upstream primer P19A1 (5'-GGTCAAGGAACACAAGATGG-3') and the downstream primer P19A2 (5'-**ATGGTGATGGTGATGGTG**TTCCAGACACCTTCTGA-3'). The latter introduced a nucleotide sequence encoding six histidine residues (bold). The PCR 1 product was the template for the second amplification using the forward primer P19B1 (5'-TCCCCGCGGAGATGGTTTTGGAAATGCTGAA-3') and reverse primer P19B2 (5'-TCCCCGCGG*TTA***ATGGTGATGGTGATGGTG**TT-3'). Both primers introduced *Sac*II sites (underlined) and the latter primer inserts a stop codon (italics) after the six histidine tag coding region (bold).

Each PCR amplification mixture (100 µl) contained pCMV-CYP19 plasmid DNA (750 ng) and the reaction was carried out as described in section 1.1.1.2. The PCR 1 product was electrophoresed on an 0.8% low melting agarose, extracted and purified as described in section 1.1.1.2. The same reaction procedure was followed for a second amplification using the purified PCR 1 product (400 ng) as a template. The PCR 2 product was electrophoresed, extracted and purified as described in section 1.1.1.2.

The CYP19 PCR 2 product was digested with *Sac*II and cloned into the pGEM-T®Easy (pGEM-T) vector (Promega) linearised with *Sac*II as depicted in Figure 1.31. Two separate restriction digestions, one containing amplified CYP19 (2 µg) and the other containing pGEM-T DNA (2 µg) were carried out at 37°C for 120 min, each in a final

---

[2] A kind gift from Prof Ian Mason, Clinical Biochemistry Section University of Edinburgh School of Clinical Sciences & Community Health, Edinburgh, Scotland, UK.

Stellenbosch University http://scholar.sun.ac.za

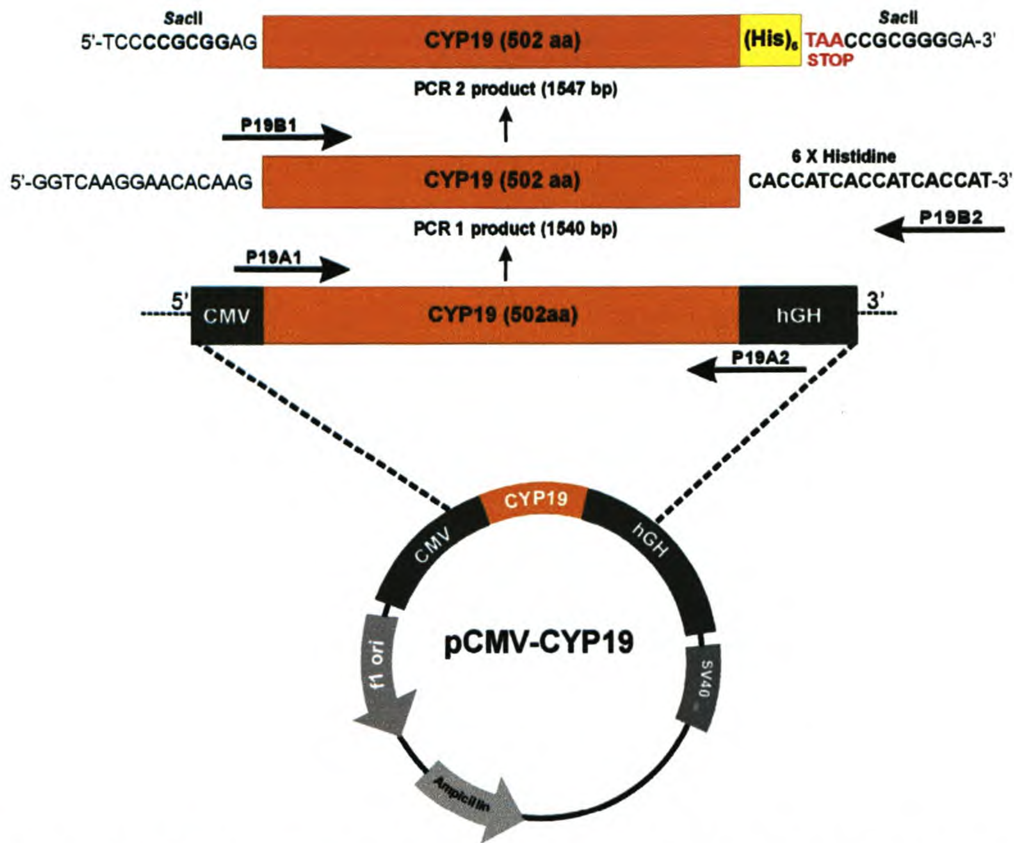

Figure 1.30 Schematic representation of the cloning strategy for the CYP19 amplified from the pCMV-CYP19 expression vector. The first PCR amplification using the upstream primer P19A1 and the downstream primer P19A2 introduced a nucleotide sequence encoding the six histidine tag at the 3'-terminal of the cDNA. Primers P19B1 and P19B2 were used in the second amplification. Both primers introduced *SacII* sites and the latter primer additionally introduced nucleotide bases encoding a stop codon after the residues encoding the six histidine tag.

volume of 30 μl containing 5.0 μl of 10 × restriction buffer D (supplied), acetylated BSA (5.0 μg) and *Sac*II (10 U). The digested CYP19 fragment and linearised vector were electrophoresed on an 0.8% low-melting agarose gel in TAE buffer (40 mM), extracted and purified as described in section 1.1.1.2.1.

A dephosphorylation reaction containing linearised pGEM-T vector (4.0 μg) was carried out at 37°C for 15 min in a final volume of 50 μl containing dephosphorylation buffer (50 mM Tris-HCl, pH 8.5, 5 mM $MgCl_2$) and shrimp alkaline phosphatase/SAP (4.0 U)(Roche). The reaction was terminated at 65°C for 15 min and the DNA was purified with the Wizard™ DNA Clean-up® kit. The purified digested CYP19 fragment (100 ng) and pGEM-T DNA (200 ng) were subsequently ligated as described in section 1.1.1.2.4. Competent JM109 cells were transformed with the ligated vectors. In addition to

Stellenbosch University http://scholar.sun.ac.za

ampicillin selection, the vector contained the lacZ gene, which allowed for blue/white colony selection.

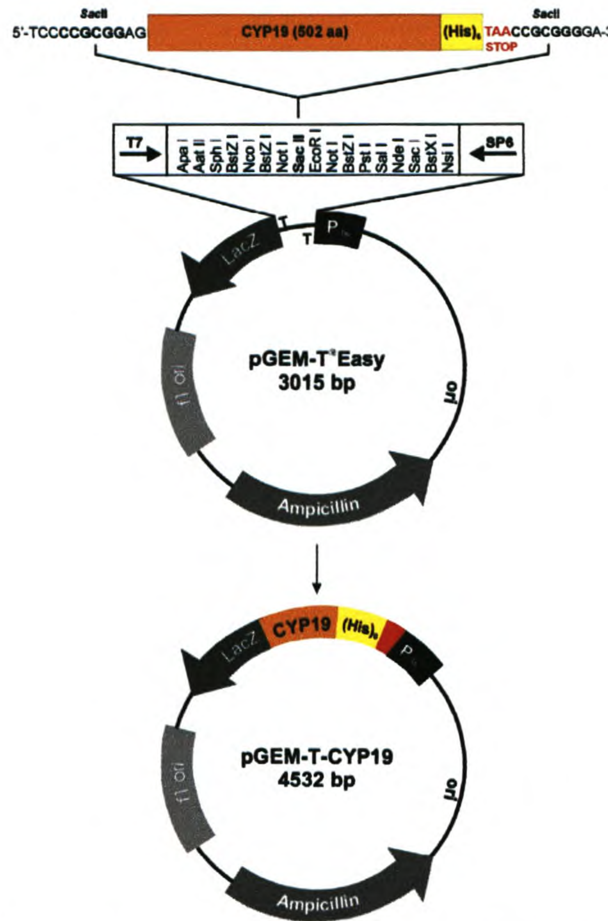

Figure 1.31 Schematic representation of the subcloning of the CYP19 into pGEM-T.

Positive colonies were subsequently incubated overnight in LB medium (5 ml) at 37°C and the plasmid DNA was isolated using the Wizard™ *Plus* Minipreps DNA Purification System as described in section 1.1.1.2.5. The purified recombinant plasmid DNA was digested with *Sac*I and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit as described in section 1.1.1.2.5. The purified isolated plasmid DNA was analysed by *Sac*I restriction digestion and DNA sequencing and subsequently used in construction of the pPIC9K-CYP19 expression vector.

Stellenbosch University http://scholar.sun.ac.za

1.4.1.2 Construction of the pPIC9K-CYP19 expression vector

The pGEM-T-CYP19 construct (4532 bp) was digested by *Not*I and the CYP19 was ligated into pPIC9K vector, previously linearised with *Not*I as depicted in Figure 1.32.

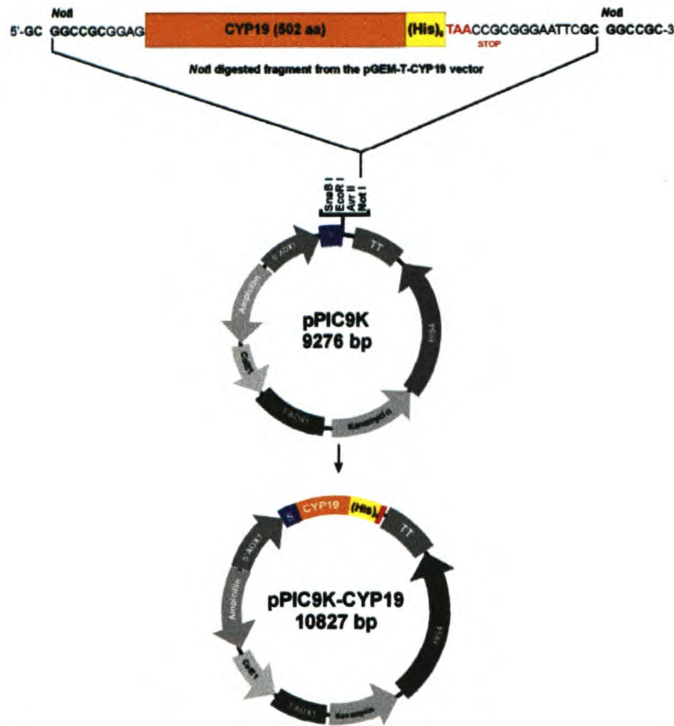

Figure 1.32 Schematic representation of the cloning strategy for the construction of the pGEM-T-CYP19 expression vector. The cloned CYP19 in the pGEM-T vector was inserted into the pPIC9K expression vector with compatible *Not*I sites.

Two separate restriction digestions, one containing pGEM-T-CYP19 DNA (1 µg) and the other containing pPIC9K DNA (1 µg) were carried out at 37°C for 90 min each in a final volume of 25 µl containing 2.5 µl SuRE/Cut buffer H (supplied) and *Not*I (2 U) restriction enzyme. The digested pPIC9K DNA was purified with the Wizard™ DNA Clean-up® kit according to the manufacturer's instructions. The digested CYP19 fragment was electrophoresed on an 0.8% low-melting agarose gel, extracted and purified using the QIAquick Gel Extraction kit (Qiagen), according to the manufacturer's instructions.

The linearised pPIC9K DNA (6 µg) was dephosphorylated with CIAP (1.96 U) and purified as described in section 1.1.1.2.3. The purified digested CYP19 (32.2 ng) and dephosphorylated linear pPIC9K DNA (200 ng) were used in a 10 µl ligation reaction

Stellenbosch University http://scholar.sun.ac.za

containing T4 ligase (1 U) (Roche), ligase buffer (66 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 1 mM DTT and 1 mM ATP) and ATP (1 mM). The reaction was incubated at 16°C for 22 hours and stored at –20°C. Competent JM109 cells were transformed with the ligated vectors and transformants were selected on LB-agar plates with ampicillin.

Positive colonies were subsequently incubated overnight in LB medium (8 ml) at 37°C and the plasmid DNA was isolated using the Wizard™ SV*Plus* Minipreps DNA Purification System (Promega), according to the manufacturer's instructions. The purified recombinant plasmids were digested with *Sac*I and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit as described in section 1.1.1.2.5.

The purified recombinant pPIC9K-CYP19 construct was analysed by DNA sequencing as described in section 1.1.1.4.3. The full-length CYP19 with its insertion sites was sequenced with seven primers as illustrated in Figure 1.33. The upstream primers were 5'AOX1P, AromB1 (5'-GTCACAGTCTGTGCTGAATC-3') and AromC1 (5'-GGTGACCTGACAAGAGAGAA-3') and the downstream primers were 3'AOX1P, AromA2 (5'-CTCCAACCTGTCCAGATGTG-3'), AromB2 (5'-GCACTGGTTCA CATTCTCTC-3') and AromC2 (5'-CACCTGT CTGAGTTTCTTGG-3').

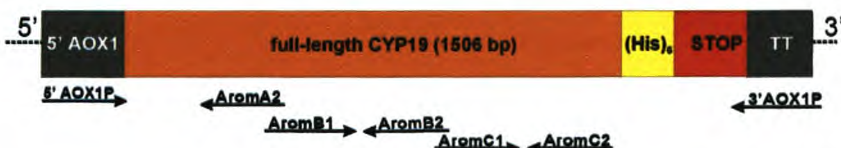

Figure 1.33 Schematic representation of the DNA sequencing strategy for the full-length human CYP19 in the pPIC9K-CYP19 expression vector.

1.4.1.3 Screening of GS115/pPIC9K-CYP19 and KM71/pPIC9K-CYP19 transformants for integration

The cloned expression vectors, pPIC9K-CYP19 (10 μg) and the parent vector pPIC9K (10 μg), were linearised with *Sal*I and transformed into competent *P. pastoris* strains, GS115 Mut^+ (Methanol utilisation plus) and KM71 Mut^s (Methanol utilisation slow)

Stellenbosch University http://scholar.sun.ac.za

using the spheroplast method as described section 1.1.1.5. The presence of CYP19 cDNA integrated into the yeast chromosomes was subsequently confirmed by PCR amplification analyses as described in section 1.1.1.5.

1.4.2 Results

1.4.2.1 Subcloning of the CYP19 cDNA

The full-length CYP19, located in the pCMV-CYP19 expression vector was used as the template to amplify a full-length CYP19 cDNA, yielding a PCR 1 product (1540 bp). This amplified product was used as a DNA template for the second amplification, which yielded the expected PCR 2 product (1547 bp)(results not shown). The amplified CYP19 was subcloned into an intermediate vector, pGEM-T, before cloning it into the pPIC9K expression vector.

The integrity of the newly constructed pGEM-T-CYP19 vector was confirmed by restriction digest analyses using *Sac*I and 1% agarose gel electrophoresis. The plasmid (4532 bp) carrying the insert yielded the expected two bands corresponding to 3378 and 1154 bp for the *Sac*I digestion as shown in Figure 1.34, lanes 1, 3, 7, 8 and 16.

Figure 1.34 Restriction digest analysis of pGEM-T plasmids containing the CYP19. Lanes 1, 3, 7, 8 and 16: *Sac*I digest of plasmid DNA (0.5 μg) containing the CYP19 in the 5'-3' orientation yielding 3378 bp and 1154 bp fragments. **Lanes 2, 6, 10, 11, 13, 14, 15, 17 and 18:** *Sac*I digest of plasmid DNA (0.5 μg) containing the CYP19 in the 3'-5' orientation. **Lanes 4 and 19:** DNA marker λDNA/*Eco*RI/*Hind*III. **Lanes 5 and 9:** Religated pGEM-T vector (0.5 μg) without an insert. **Lane 12:** DNA molecular weight marker VI.

Stellenbosch University http://scholar.sun.ac.za

1.4.2.2 Construction of the pPIC9K-CYP19 expression vector

The *Not* I digested pGEM-T-CYP19 and pPIC9K plasmid DNA were electrophoresed on an 0.8% low melting agarose gel and the CYP19 and linearised plasmid DNA were excised and purified. Prior to ligation the cohesive compatible ends of the *Not* I linearised pGEM-T-CYP19 DNA were dephosphorylated and the CYP19 successfully ligated.

Eighteen positively transformed bacterial colonies were grown in small-scale cultures. Plasmid DNA isolated from each culture was digested with *Sac*I and analysed by 1% agarose gel electrophoresis. The cloned pGEM-T-CYP19 vector (10827 bp) yielded the correct 9371 and 1456 bp fragments as seen in Figure 1.35, lanes 12 and 14.

DNA sequencing analyses, Appendix A.4, of these clones confirmed that the CYP19 had been inserted in reading frame for eukaryotic translation initiation and secretion.

Figure 1.35 Restriction digest analysis of pPIC9K plasmids containing CYP19. Lane 1: Positive control containing *Sac*I digested pPIC9K plasmid DNA (0.5 μg). **Lanes 2-4, 7, 15 and 17-20:** *Sac*I digest of plasmid DNA (0.5 μg) containing the CYP19 in the 5'-3' orientation yielding 9371 and 1456 bp fragments. **Lanes 5, 6, 8, 10-14, 16 and 17:** *Sac*I digest of plasmid DNA (0.5 μg) containing the insert in the 3'-5' orientation. **Lane 9:** DNA marker λDNA/*Eco*RI/*Hind*III.

1.4.2.3 Screening of GS115/pPIC9K-CYP19 and KM71/pPIC9K-CYP19 transformants for integration

PCR followed by 1% agarose gel electrophoretic analyses of lysed His$^+$ transformants could not determine if the CYP19 was integrated correctly, since the expected 2 bands, have similar sizes and could not be separated. One corresponded to the CYP19 (2044

Stellenbosch University http://scholar.sun.ac.za

bp) and the other to the *AOX1* gene (2100 bp) as shown in Figure 1.36, lane 3. A gene-specific primer AromB1 together with 5'AOX1P primer was used to confirm the presence of CYP19 (1171 bp) as shown in Figure 1.36, lane 4. Mut^{+} integrants with only the parent vector pPIC9K produced 492 bp and 2100 bp PCR products seen in Figure 1.36, lane 5.

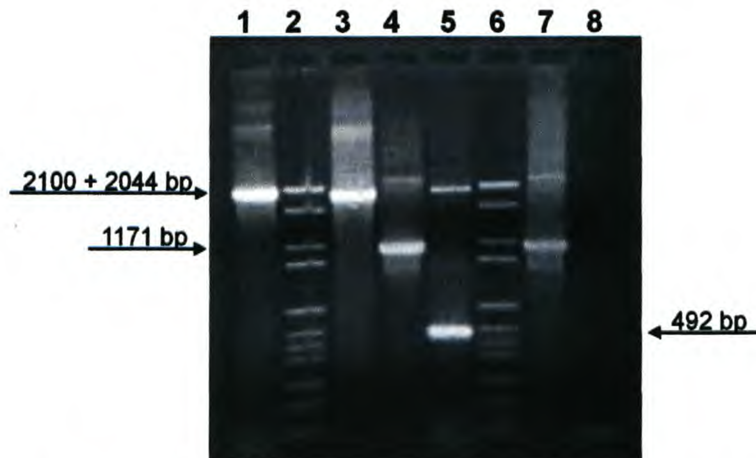

Figure 1.36 Direct PCR screening of GS115/pPIC9K-CYP19 clones. Lane 1: Amplification of the wild-type *AOX1* gene (2100 bp) using AOX1P primers. **Lanes 2 and 6:** DNA molecular weight marker VI. **Lanes 3 and 4:** Contains a clone carrying both the *AOX1* gene and CYP19. In lane 3, AOX1P primers were used resulting in the 2100 bp product containing the *AOX1* gene (2100 bp) and the 2044 bp product containing the CYP19 (1552 bp) together with 492 bp of flanking *AOX1* sequence from pPIC9K-CYP19. In lane 4, insert-specific primer AromB1 and the 3'AOX1P primer were used resulting in a 1171 bp PCR product. **Lane 5:** Contains a clone with the *AOX1* gene (2100 bp) carrying the pPIC9K vector without insert (492 bp). AOX1P primers were used. **Lane 7:** Positive PCR control containing CYP19 fragment (1171 bp) amplified from pPIC9K-CY19 plasmid using AOX1P primers. **Lane 8:** Negative control, PCR amplification without DNA template. 20 µl reaction mixtures were loaded.

The successful integration of the CYP19 into the yeast genomic DNA was shown by *Sac*I restriction digestion of the amplified PCR products. The integrated CYP19 with one *Sac*I restriction site, produced three fragments. Two corresponded to the CYP19 (1233 and 811 bp) and the remaining to the *AOX1* gene (2100 bp) as shown in Figure 1.37, lane 4. The same clone was also digested with *Xba*I. The integrated CYP19 has no *SacI* restriction site, but the *AOX1* gene has one site, which resulted in three bands. Two corresponded to the AOXI (1233 and 811 bp) and the remaining to the CYP19 (2044 bp) as shown in Figure 1.37, lane 6.

The same PCR analysis was carried out on KM71 integrants. The results were similar, except that the wild-type *AOX1* gene in KM71 cells was approximately 4500 bp instead

Stellenbosch University http://scholar.sun.ac.za

of 2100 bp, because the *ARG4* gene was inserted into the *AOX1* gene. Two bands were produced. One corresponded to the CYP19 (2044 bp) and the remaining to the *AOX1* gene (4500 bp) as shown in Figure 1.38, lane 3.

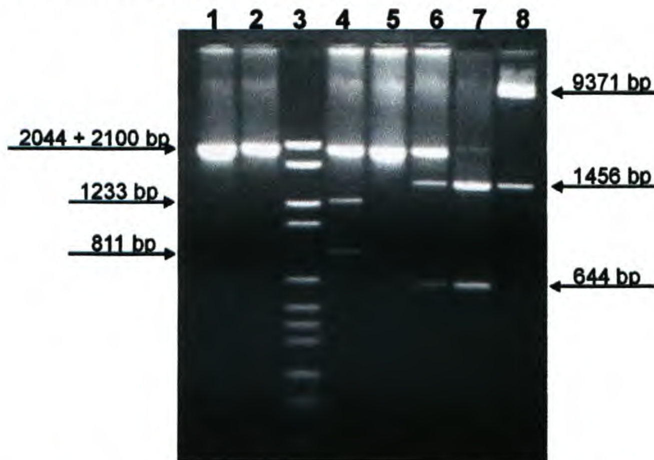

Figure 1.37 Restriction digest analysis of the PCR products from GS115/pPIC9K-CYP19 clones. The reaction products (0.5 μg) were analysed by 1% agarose gel electrophoresis. **Lanes 1 and 2:** undigested and *Sac*I digested PCR products of the wild type *AOX1* gene), respectively. **Lane 3:** DNA molecular weight marker VI. **Lane 4:** *Sac*I digested PCR products of the clone carrying both the *AOX1* gene and the CYP19. The CYP19 produced two bands (1233 and 811 bp) and the *AOX1* gene produced the remaining band (2100 bp). **Lanes 5 and 6**: undigested and *Xba*I digested PCR product. The CYP19 produced an undigested band (2044 bp) and the *AOX1* gene produced the remaining two bands (1456 and 644 bp). **Lane 7:** *Xba*I digested PCR product of the *AOX1* gene resulted in two bands (1456 and 644 bp). **Lane 8:** Positive control containing *Sac*I digested pPIC9K-CYP19 plasmid yielding bands 9371 and 1456 bp.

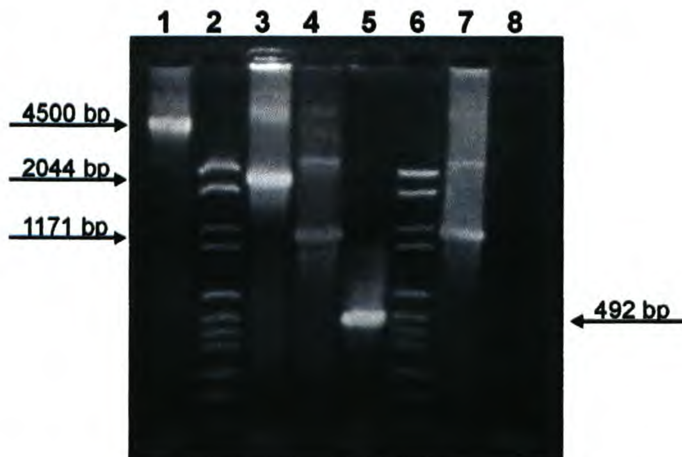

Figure 1.38 Direct PCR screening of KM71/pPIC9K-CYP19 clones. The reaction products (20 μl) were analysed by 1% agarose gel electrophoresis. **Lane 1:** Amplification of the wild-type *AOX1* gene with *ARG4* insert (4500 bp) using AOX1P primers. **Lanes 2 and 6:** DNA molecular weight marker VI. **Lanes 3 and 4:** Contains a clone carrying both the *AOX1* gene (4500 bp) and CYP19. In lane 3, AOX1P primers were used resulting in two PCR products. One corresponded to the CYP19 (2044 bp) and the other to the *AOX1* gene (4500 bp). In lane 4 insert-specific primer AromB1 and the 3'AOX1P primer were used resulting in an 1171 bp PCR product. **Lane 5:** Contains a clone carrying the parent pPIC9K vector without insert (492 bp). AOX1P primers were used. **Lane 7:** Positive PCR control containing CYP19 fragment (1171 bp) amplified from pPIC9K-CY19 plasmid using AOX1P primers. **Lane 8:** Negative control, PCR amplification without DNA template.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 2 : CULTURE OF *P. PASTORIS* EXPRESSING DIFFERENT PROTEINS

2.1 Growth and expression of P45017α, aromatase and b_5 in *P. pastoris*

2.1.1 Materials and Methods

2.1.1.1 Growth

Clones that had been verified by direct PCR were tested for the expression of the desired protein. Single colonies of *P. pastoris* strains (GS115 and KM71) transformed with pPIC9K-CYP17, pPIC9K-truncCYP17, pPIC9K-CYB5 and pPIC9K-CYP19 expression vectors were grown for in BMGY (25 ml) to an optical density of 600 nm (OD_{600}) between 2 and 6 at 275 rpm for 20 hours at 30°C in an Innova 4000 shaking incubator (New Brunswick). Yeast cells were harvested by centrifugation at 2500 × g for 5 min at room temperature and resuspended to approximately OD_{600} = 1 in BMMY induction media (200 ml). Methanol-induced cultures were grown at 30°C with shaking at 275 rpm. Methanol was added to 0.5% (v/v) final concentration after every 24 h and the induction was carried out up to 96 hours. To determine the optimal time for expression, aliquots (30 ml) of the culture media were removed at selected time intervals i.e. 0, 24, 48, 72 and 96 hours. The expression culture was centrifuged and the cells were separated from the induction media. The supernatant and pellet, resuspended in buffer BB (50 mM sodium phosphate, pH 7.6, 1 mM EDTA, 5% glycerol and 1 mM PMSF), were stored at –80°C. Both fractions were analysed for the presence of extracellularly expressed and intracellular heterologous P450 proteins by SDS-PAGE (Coomassie staining) and immunoblot analyses.

2.1.1.2 Screening of yeasts for secretory expression of proteins

Culture media aliquots (30 ml) were removed at specific time intervals and stored at –80°C. The frozen samples were defrosted on ice and concentrated to a predetermined volume as described by Swart *et al.* (57). The dialysis tube used was Spectapor No. 1 (20.4 mm inflated diameter, 6000-8000 MCO) from Spectrapor, LA and the

Stellenbosch University http://scholar.sun.ac.za

concentration was run in polyethylene glycol 35.000 (Fluka) (20% w/v) overnight at 4°C. The concentrated suspension (4 ml) was centrifuged at 12000 × g. at 4°C for 10 min and the clear media was further concentrated with an Ultrafree-15 Centrifugal filter device (Millipore), according to the manufacturer's instructions. The concentrated medium was recovered from the device and stored at –20°C for SDS-PAGE and immunoblot analyses.

2.1.1.3 Screening for intracellular expression of the proteins

A rapid procedure was used to obtain a crude membrane fraction from cells grown by batch cultivation, to monitor specific P450 protein expression. Cell pellets, stored at –80°C, were thawed on ice and resuspended in breaking buffer (BB) (1 ml)(50 mM NaH_2PO_4, pH 7.4, 1 mM PMSF, 1 mM EDTA and 5% glycerol). All further steps were carried out at 4°C. The suspension was transferred to a 2 ml microcentrifuge tube and centrifuged at 3000 × g at 4°C for 7 min. The washed cells were resuspended in BB (0.5 ml, OD_{600} of 100-125) and an equal amount of acid washed glass beads (Sigma G-9268, 425-600 μm) was added. The mixture was vortexed vigorously 10 times for 30 s each, with 30 s intervals on ice. The supernatant was centrifuged at 12000 × g for 10 min to remove the glass beads and cell debris. The clear lysate (200-250 μl) was pipetted off carefully and stored at –20°C.

2.1.1.4 SDS-PAGE and Western blot analyses

SDS-PAGE was carried out according to the method of Laemmli (58) using an acrylamide concentration of 10% (v/v), except for the b_5 expression experiments, which used 12% SDS-PAGE. Low Range molecular weight markers (Biorad) and Rainbow coloured protein molecular weight markers (Amersham Pharmacia) consisting of myosin (220 kDa), phosphorylase b (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (21 kDa) and lysozyme (14.4 kDa) were used. Wide range ColorBurst® Electrophoresis Markers (Sigma) were used alternatively to the Rainbow Markers, which have a 60 kDa protein marker instead of 66 kDa. Proteins were stained with Coomassie brilliant blue.

Stellenbosch University http://scholar.sun.ac.za

Electroblotting was carried out in a wet blotting system. Proteins were transferred onto a nitrocellulose membrane (Schleicher&Schuel) and the antigen-probe complex, which is a nickel (Ni^{2+}) activated horseradish-peroxidase (INDIA™ HisProbe-HRP, Pierce, Rockford, USA) was visualised by chemiluminescence (SuperSignal West HisProbe Kit, Pierce, Rockford. IL). The probe was diluted 1:2000 in TBST buffer. A six histidine tagged ureate oxidase (Pierce, Rockford, IL) was used as a positive control.

2.1.2 Results

2.1.2.1 **Growth**

GS115 transformants (Mut^{+}) were grown in BMGY media for 18 hours in the log-phase. They were harvested and resuspended in expression media (BMMY) at ~ OD_{600} of 1 and grown for four days supplementing with methanol every 24 hours. The cell density of cultures during the induction phase increased linearly up to 48 hours after which it decreased into stationary phase as shown in Figure 2.1. The growth of the transformants was similar to the growth of the control strain GS115/β-Gal (Mut^{+}). The wild-type strain GS115 had a similar initial growth rate as the cloned transformants, but after 24 hours the rate was maintained to 48 hours doubling the cell density before decreasing and changing into stationary phase. This high growth rate was also observed with the GS115/pPIC9K-CYB5 clone.

KM71 transformants (Mut^{s}) were grown and induced the same way as the GS115 strains. The cultures grew in a linear fashion up to 48 hours, but then abruptly reached stationary phase as shown in Figure 2.2.

All the GS115 and KM71 strains, transformed with pPIC9K-CYP17, pPIC9K-truncCYP17, pPIC9K-CYB5 and pPIC9K-CYP19 constructs, were characterised by Western blot analyses. The supernatant (induction media) was concentrated and analysed for extracellular expression. Low levels or no external detection of the expressed hemo-proteins were found and the cells were subsequently analysed for intracellular localisation of expressed proteins.

Stellenbosch University http://scholar.sun.ac.za

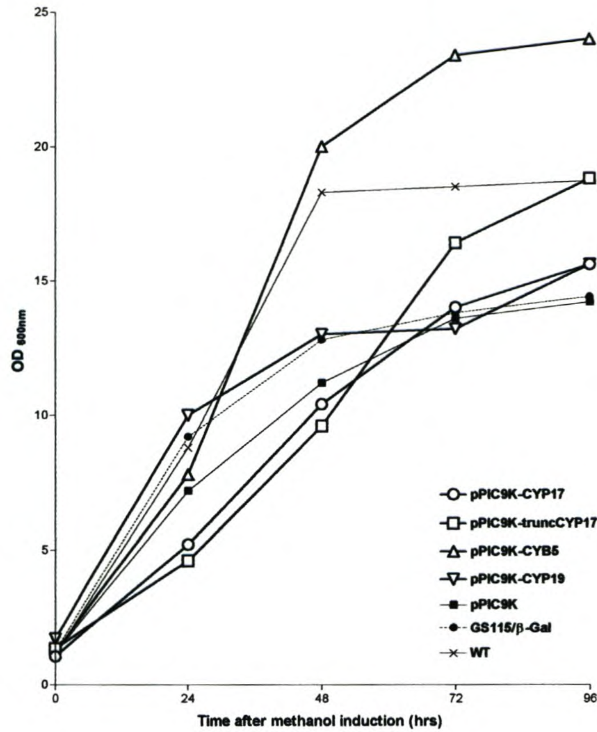

Figure 2.1 Growth curve of GS115 transformants during induction phase. Growth of the strains was monitored by optical density at 600 nm after the addition of methanol (0.5%, v/v final concentration). The medium was supplemented with methanol at 24 hours intervals. GS115/β-Gal is a Mut^+ control strain and WT is the wild-type GS115 strain.

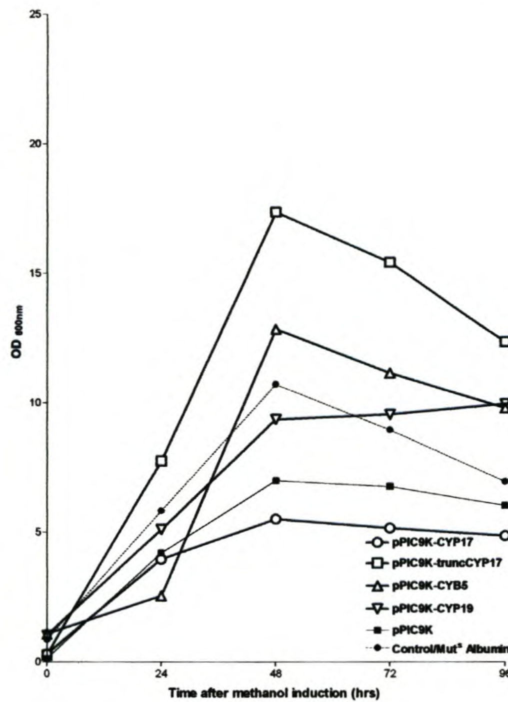

Figure 2.2 Growth curve of KM71 transformants during induction phase. Growth of strains was monitored at OD_{600} after the addition of methanol (0.5%, v/v final concentration). The medium was supplemented with methanol at 24 hours intervals. Control/Albumin is a Mut^s control strain.

Stellenbosch University http://scholar.sun.ac.za

2.1.2.2 Expression of full-length and truncated P45017α by GS115 cells

Secretory expression of the full-length P45017α in the induction media (30 times concentrated), analysed by SDS-PAGE and Western blot analyses, was slightly detectable at 96 hours as seen in Figure 2.3B, lane 8. GS115 transformed with the parent vector pPIC9K without an insert was used as the negative control and a background protein band of 58 kDa was observed, Figure 2.3, lanes 1-2.

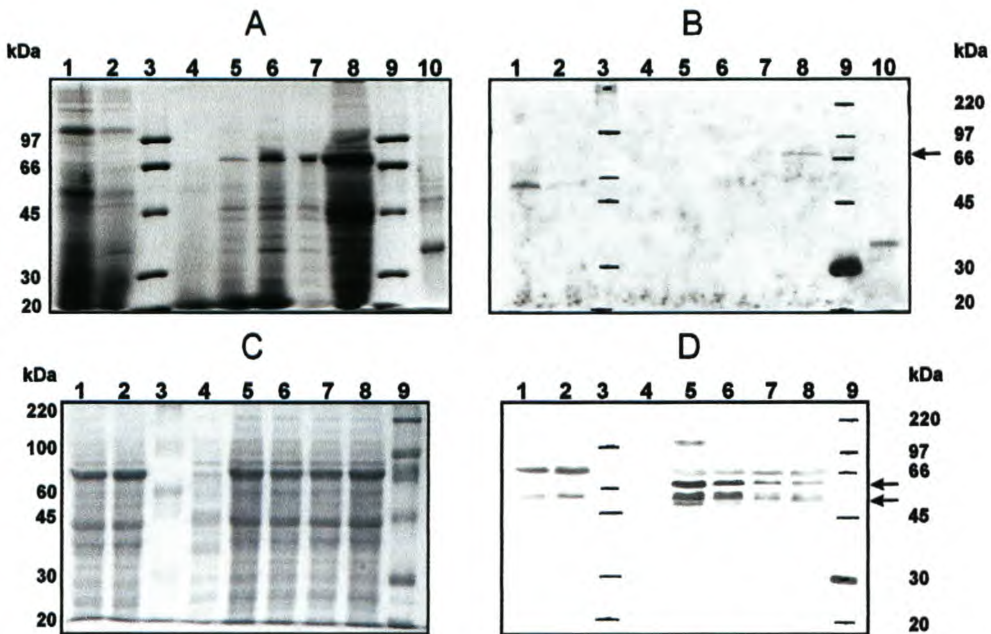

Figure 2.3 SDS-PAGE (A) and Western blot (B) analyses of proteins from the induction media and SDS-PAGE (C) and Western blot (D) analyses of proteins in the intracellular fraction of GS115 cells transformed with pPIC9K-CYP17 expression vector. Lanes 1 and 2: Negative control containing total proteins of GS115/pPIC9K cell cultures at 72 and 96 h, respectively. **Lanes 3 and 9:** Low-molecular weight Marker. **Lanes 4-8:** Total proteins of GS115/pPIC9K-CYP17 cell cultures at 0, 24, 48, 72 and 96 hours. **Lane 10:** Positive control containing histidine tagged ureate oxidase. Total protein concentration (10 μg) per well was loaded for the intracellular fractions. The arrow indicates a 62 kDa and 55 kDa histidine tagged P45017α protein. Detection was carried out using an INDIA HisProbe-HRP at 1:2000.

Immunoblot analysis of induction media of GS115, transformed with pPIC9K-truncCYP17 expression vector showed similar results. Secretory expression of histidine tagged truncated P45017α was detectable at 96 hours (Figure 2.4B, lanes 4-7). However, four faint bands were observed, as seen in Figure 2.4B, lane 8. GS115 strain transformed with the parent vector pPIC9K without an insert was used as the negative control (Figure 2.4, lanes 1-2).

Stellenbosch University http://scholar.sun.ac.za

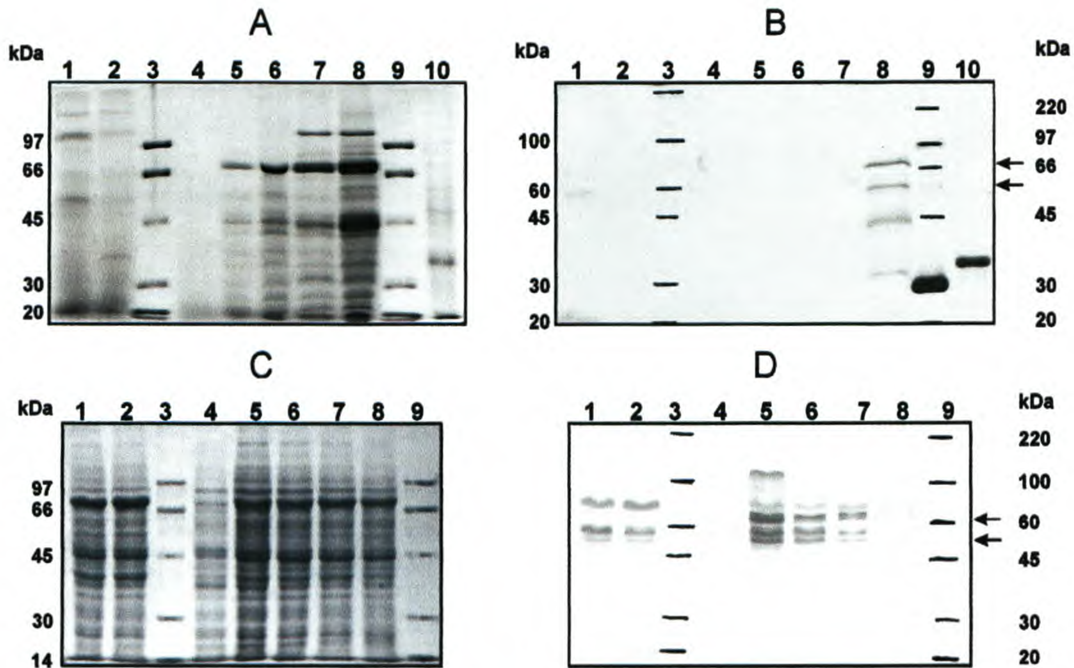

Figure 2.4 SDS-PAGE (A) and Western blot (B) analyses of proteins from the induction media and SDS-PAGE (C) and Western blot (D) analyses of proteins in the intracellular fraction of GS115 cells transformed with pPIC9K-truncCYP17 expression vector. Lanes 1 and 2: Negative control containing total proteins of GS115/pPIC9K cell cultures at 72 and 96 h, respectively. **Lanes 3 and 9:** Low-molecular weight Marker. **Lanes 4-8:** Total proteins of GS115/pPIC9K-truncCYP17 cell cultures at 0, 24, 48, 72 and 96 hours. **Lane 10:** Positive control containing histidine tagged ureate oxidase. Total protein concentration (10 μg) per well was loaded for the intracellular fractions. The arrow indicates a 62 kDa and 55 kDa histidine tagged P45017α protein. Detection was carried out using an INDIA HisProbe-HRP at 1:2000.

The GS115/pPIC9K-CYP17 and GS115/pPIC9K-truncCYP17 transformants at different time course were lysed and the intracellular fractions analysed by SDS-PAGE and Western blot analyses. The lysed GS115 cells contained a histidine tagged P45017α protein of an estimated molecular weight of 62 kDa (Figure 2.3D, lanes 5-8. The total protein concentration of the samples loaded per well was kept equal and the histidine signal, hence P45017α concentration was the highest at 24 hours and decreased with time. Western blot analysis of the intracellular fraction of the control strain GS115/pPIC9K as shown in Figure 2.3D, lanes 1-2, has two background proteins, size 58 and 66 kDa. Similar results were observed from lysed GS115/pPIC9K-truncCYP17 cells expressing the truncated P45017α (62 kDa) as seen in Figure 2.4 D, lanes 5-8.

Stellenbosch University http://scholar.sun.ac.za

2.1.2.3 Expression of full-length P45017α by *P. pastoris* KM71

Secretory expression of the full-length P45017α in the induction media (30 times concentrated), analysed by SDS-PAGE and Western blot analyses, was not detectable Figure 2.5B.

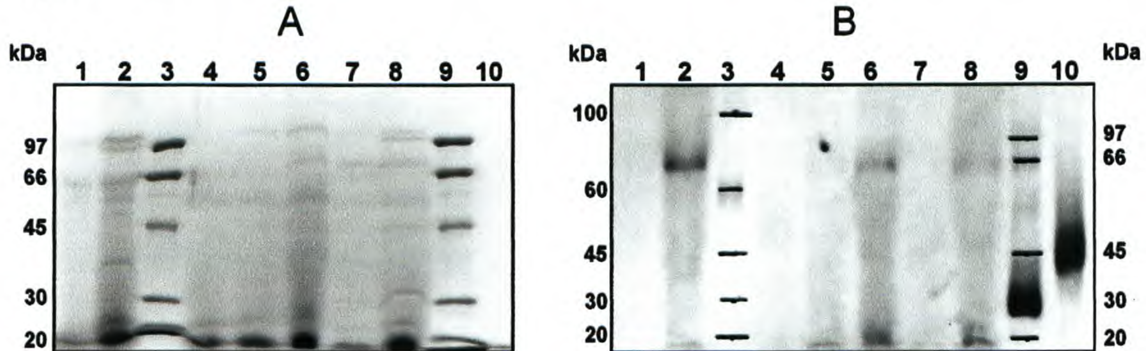

Figure 2.5 SDS-PAGE (A) and Western blot (B) analyses of proteins from the induction media of KM71 cells transformed with pPIC9K-CYP17 expression vector. Lanes 1 and 2: Negative control containing total extracellular proteins of KM71/pPIC9K cell cultures at 72 and 96 h, respectively. **Lanes 3 and 9:** Low-molecular weight Marker. **Lanes 4-8:** Total extracellular proteins (10 µg) of KM71/pPIC9K-CYP17 cell cultures at 0, 24, 48, 72 and 96 hours. **Lane 10:** Positive control containing histidine tagged ureate oxidase. Detection was carried out using an INDIA HisProbe-HRP at 1:2000.

2.1.2.4 Expression of b_5 by *P. pastoris* GS115

Secretory expression of the full-length b_5 in the induction media (100 times concentrated), analysed by SDS-PAGE and Western blot analyses, was detectable at 48 and 72 hours as seen in Figure 2.6B, lanes 6 and 7. However b_5 has a molecular weight of 16 kDa and the estimated 64 kDa protein band might be an aggregated tetramer (59). GS115 transformed with the parent vector pPIC9K without an insert was used as the negative control (Figure 2.6B, lanes 1-2).

The GS115/pPIC9K-CYB5 transformants at different time course were lysed and the intracellular fractions analysed by Western blot analyses. The lysed GS115 cells contained no detectable histidine tagged b_5 protein (Figure 2.6D, lanes 4-8).

Stellenbosch University http://scholar.sun.ac.za

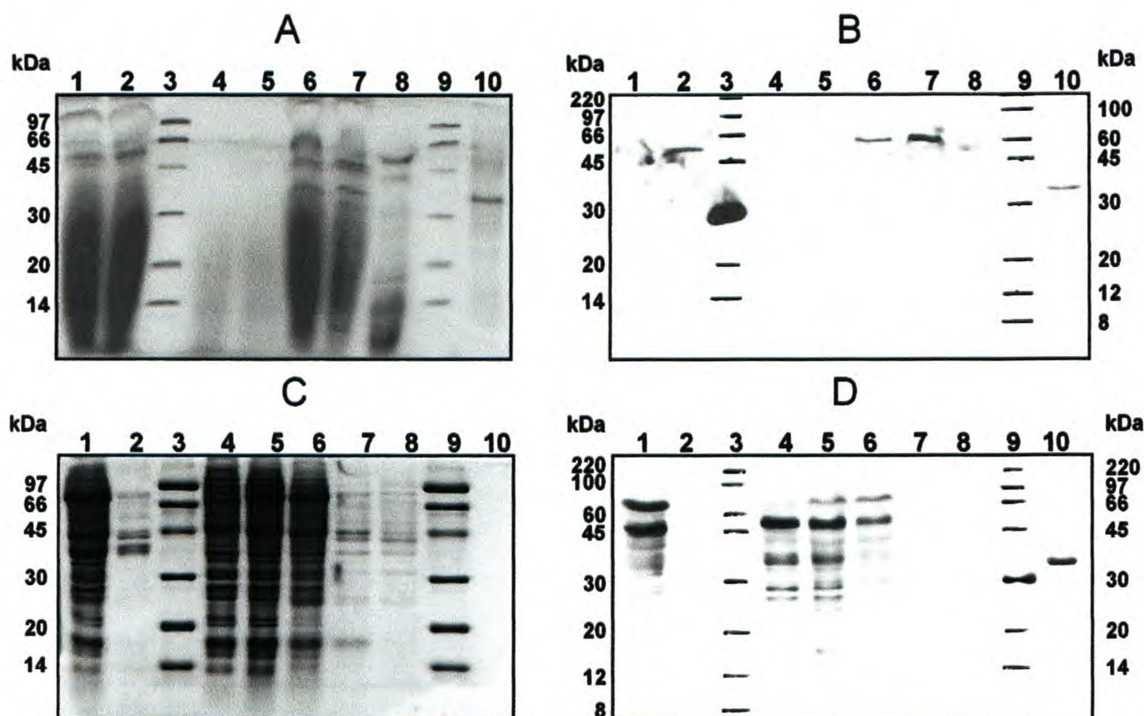

Figure 2.6 SDS-PAGE (A) and Western blot (B) analyses of proteins from the induction media and SDS-PAGE (C) and Western blot (D) analyses of proteins in the intracellular fraction of GS115 cells transformed with pPIC9K-CYB5 expression vector. Lanes 1 and 2: Negative control containing total proteins of GS115/pPIC9K cell cultures at 72 and 96 h, respectively. **Lanes 3 and 9:** Low-molecular weight Marker. **Lanes 4-8:** Total proteins of GS115/pPIC9K-CYB5 cell cultures at 0, 24, 48, 72 and 96 hours. **Lane 10:** Positive control containing histidine tagged ureate oxidase. Total protein concentration (10 μg) per well was loaded for the intracellular fractions (lane 1 and 4-6). Detection was carried out using an INDIA HisProbe-HRP at 1:2000.

2.1.2.5 Expression of b_5 by *P. pastoris* KM71

Secretory expression of the full-length b_5 in the induction media (100 times concentrated), analysed by SDS-PAGE and Western blot analyses, was undetectable using polyclonal mouse anti-sheep b_5 antibodies (see Figure 2.7B, lanes 4-8) or using INDIA HisProbe-HRP (Figure 2.7C, lanes 4-8) at 48 and 72 hours as seen in Figure 2.6B, lanes 6 and 7. However b_5 has a molecular weight of 16 kDa and the estimated 64 kDa protein band might be an aggregated tetramer (59). GS115 transformed with the parent vector pPIC9K without an insert was used as the negative control (Figure 2.6B, lanes 1-2). A faint band at approximately 58 kDa was also detected in the negative control.

Stellenbosch University http://scholar.sun.ac.za

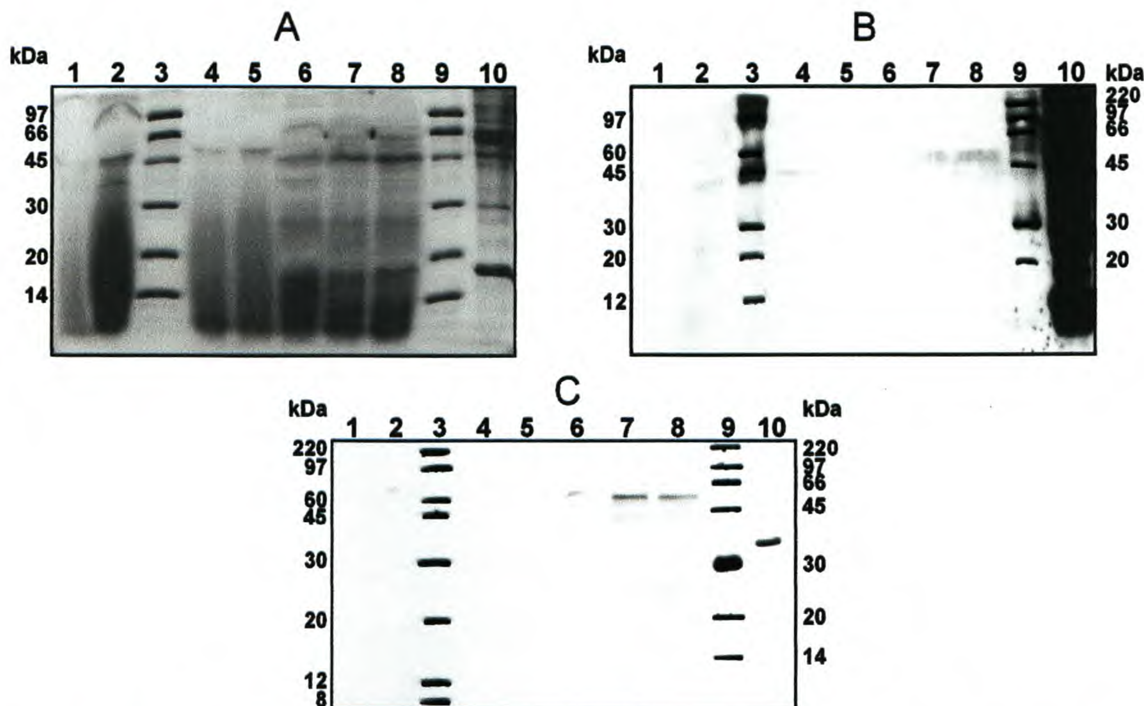

Figure 2.7 SDS-PAGE (A) and Western blot (B and C) analyses of proteins from the induction media of KM71 cells transformed with pPIC9K-CYB5 expression vector. Lanes 1 and 2: Negative control containing total induction media proteins of KM71/pPIC9K cell cultures at 72 and 96 h, respectively. **Lanes 3 and 9:** Low-molecular weight Markers. **Lanes 4-8:** Total extracellular proteins (10 μg) of KM71/pPIC9K-CYB5 cell cultures at 0, 24, 48, 72 and 96 hours. **Lane 10:** Western blot **B** contains b_5 in a crude microsomal extract as a positive control. Detection was carried out using polyclonal mouse anti-sheep b_5 antibodies at 1:2000. Western blot **C** contains histidine tagged ureate oxidase as a positive control. Detection was carried out using an INDIA HisProbe-HRP at 1:2000.

2.1.2.6 Expression of aromatase by *P. pastoris* GS115

Secretory expression of the full-length aromatase in the induction media (80 times concentrated), analysed by SDS-PAGE and Western blot analyses, was detectable at 48, 72 and 96 hours as seen in Figure 2.8B, lanes 6, 7 and 8. Two protein bands corresponding to molecular masses of 55 and 68 kDa are visible, however, the lower molecular mass protein is also present in the control strain. Therefore the higher molecular mass protein might be histidine tagged aromatase. GS115 strain transformed with the parent vector pPIC9K without an insert was used as the negative control (Figure 2.8, lanes 1-2).

The GS115/pPIC9K-CYP19 transformants at different time course were lysed and the intracellular fractions analysed by Western blot analyses. The lysed GS115 cells contained a histidine tagged aromatase protein of an estimated molecular weight of 62

Stellenbosch University http://scholar.sun.ac.za

kDa as shown in Figure 2.8D, lanes 5-8. The total protein concentration of the samples loaded per well were kept equal and the histidine signal corresponding to a molecular mass of approximately 64 kDa appeared at 72 and 96 hours as indicated by an arrow in Figure 2.8D, lanes 7 and 8.

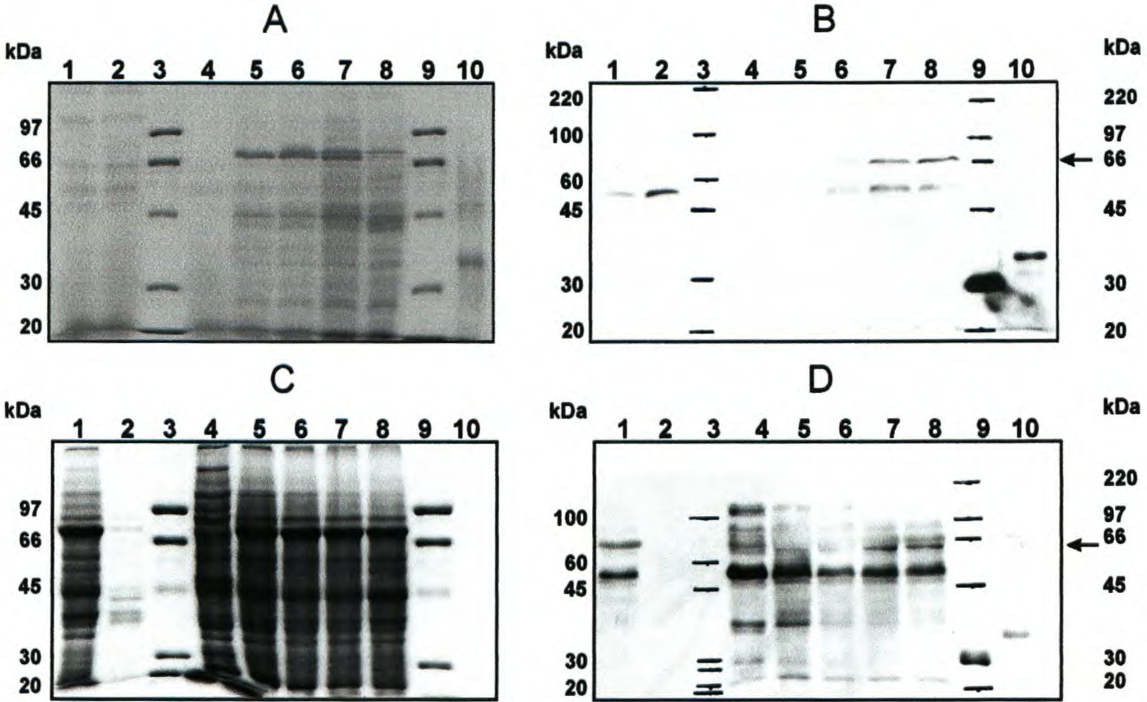

Figure 2.8 SDS-PAGE (A) and Western blot (B) analyses of proteins from the induction media and SDS-PAGE (C) and Western blot (D) analyses of proteins in the intracellular fraction of GS115 cells transformed with pPIC9K-CYP19 expression vector. Lanes 1 and 2: Negative control containing total proteins of GS115/pPIC9K cell cultures at 72 and 96 h, respectively. **Lanes 3 and 9:** Low-molecular weight Marker. **Lanes 4-8:** Total proteins (10 µg) of GS115/pPIC9K-CYP19 cell cultures at 0, 24, 48, 72 and 96 hours. **Lane 10:** Positive control containing histidine tagged ureate oxidase. The arrow indicates a 68 kDa and 64 kDa histidine tagged aromatase protein in the medium and intracellularly, respectively. Detection was carried out using an INDIA HisProbe-HRP at 1:2000.

2.1.2.7 Expression of aromatase by *P. pastoris* KM71

Secretory expression of the full-length aromatase in the induction media (30 times concentrated), analysed by SDS-PAGE and Western blot analyses, was not detectable (Figure 2.9B, lanes 4-8).

Stellenbosch University http://scholar.sun.ac.za

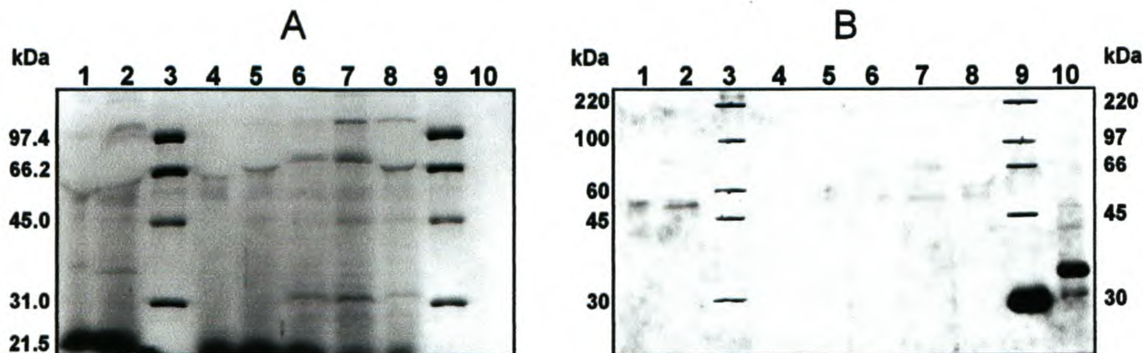

Figure 2.9 SDS-PAGE (A) and Western blot (B) analyses of proteins from the induction media of KM71 cells transformed with pPIC9K-CYP19 expression vector. Lanes 1 and 2: Negative control containing total extracellular proteins of KM71/pPIC9K cell cultures at 72 and 96 h, respectively. **Lanes 3 and 9:** Low-molecular weight Marker. **Lanes 4-8:** Total extracellular proteins (10 μg) of KM71/pPIC9K-CYP19 cell cultures at 0, 24, 48, 72 and 96 hours. **Lane 10:** Positive control containing histidine tagged ureate oxidase. Detection was carried out using an INDIA HisProbe-HRP at 1:2000.

2.1.2.8 Control strains

The GS115/His$^+$ Mut^s Albumin strain was a control for expression and secretion and the molecular mass of the protein was 67 kDa as shown in Figure 2.10.

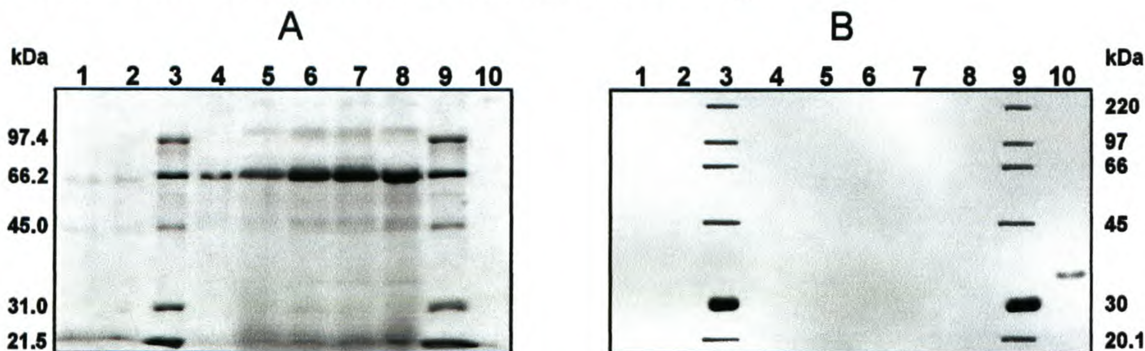

Figure 2.10 SDS-PAGE (A) and Western blot (B) analyses of albumin proteins from the induction media of *P. pastoris* GS115/His$^+$ Mut^s albumin cell culture. Lanes 1 and 2: Negative controls of *P. pastoris* with only the pPIC9K plasmid. **Lanes 3 and 9:** Rainbow Markers. **Lanes 4-8:** Expressed albumin at 6, 24, 48, 72 and 96 hours, respectively. Total extracellular proteins (10 μg) of GS115/His$^+$ Mut^s albumin cell culture at 0, 24, 48, 72 and 96 hours. **Lane 10:** Positive control containing histidine tagged ureate oxidase.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 3 : DISCUSSION

In cytochrome P450 research, the use of heterologous expression systems for the expression of cDNA encoding specific cytochromes P450 has given researchers a tool to obtain and study various, less abundant cytochromes P450. Especially human cytochromes P450, which are extremely labile and occur in low concentrations in tissues, can only be studied when heterologously expressed. Many benefits arise from using heterologous expression systems which include — (i) determination of cytochrome P450 enzymatic activities in an environment which has a low cytochrome P450 background; (ii) determination of structure/function relationships associated with cytochromes P450 by site-directed mutagenesis and (iii) evaluation of xenobiotics on the effect of expressed cytochromes. The most important aspect of heterologous expression for this project was, however, the preparation of sufficient pure P450 proteins for polyclonal and monoclonal antibody production and ultimately possible crystallisation.

Several different heterologous expression systems have been used to express P450 enzymes, including *E. coli* (60, 61, 62, 63, 64), S. cerevisiae (65), baculoviral cells (66), vaccinia virus (67), COS cells (68, 69, 70, 71), mammalian cells (72) and recently non-conventional yeast such as *P. pastoris* (73, 74, 75). Although most of these systems yielded sufficient P450 enzymes, in microsomal fractions for investigative purposes, only a few were successful in producing P450 enzymes on a relatively large-scale.

To date, all heterologously expressed P450 proteins were cytoplasmically expressed. Intracellular expression systems have the following disadvantages. Firstly, in order to achieve high-level cytochrome P450 expression, strong promoters (i.e. T7) need to be used which result in either no expression or the appearance of P450 protein in *E. coli* inclusion bodies. Secondly, it was found that the amino terminal of several cytochrome P450 enzymes needed to be modified to prevent the possibility of secondary structure formation, thereby inhibiting the initiation of translation. Thirdly, to gain access to the intracellular expressed cytochrome P450, it is necessary to disrupt the cell wall either mechanically or enzymatically, which is a tedious and timeconsuming method. Often strong denaturants are used and the yield of microsomal proteins is not very good. The expressed P450 proteins located in the microsomes need to be solubilised using strong surfactants, which could remove the heme from the protein.

Stellenbosch University http://scholar.sun.ac.za

In this study we used an extracellular expression system, to secrete human P45017α, b_5 and aromatase into the media, where it would be readily available for purification and characterisation. The cDNA of these microsomal enzymes were cloned in frame with the secretion signal open reading frame, which is a α-factor prepro peptide from *S. cerevisiae*. A nickel chelated probe was used for all the Western blot analyses, which not exclusively detects histidine tagged proteins but can also identify histidine rich proteins and metalloenzymes. Therefore the Western blot results were regarded as preliminary results and further characterisation of expressed proteins need to be investigated. Protein fractions from yeast cells transformed with the parent vector (indicating any background proteins) however, did not produce the same target proteins as seen from the recombinant P450 expressing cells.

SDS-PAGE and Western blot analyses of the concentrated cell culture medium of GS115 cells expressing the full-length P45017α enzyme only showed a faint band corresponding to a molecular mass of 68 kDa at 96 hours. However, the yeast cells did produce a histidine tagged P45017α protein, which was detected in the cytoplasmic fraction of the cells. Two distinct protein bands of 55 and 62 kDa were detected and the concentration, according to the intensity of the chemiluminescent signal, gradually decreased from 24 to 96 hours. It is most likely that the higher molecular mass protein is the uncleaved P45017α protein, which is linked to the α-factor secretion signal and the smaller molecular mass protein is the correctly processed and folded P45017α protein in the secretory pathway. The reason for the low extracellular expression of P45017α may be because the correctly translated protein is retained in the ER of the yeast cell. All microsomal cytochromes P450, including P45017α, are membrane bound proteins, which means they have a hydrophobic domain, i.e. the signal anchor sequence (SAS). It is believed that, by deleting the SAS from the amino terminal of microsomal cytochromes P450, a soluble form of P450 enzyme might be produced which could be secreted into the media unhindered by any membranes. We therefore truncated the human P45017α by deleting the first 18 hydrophobic amino acids and inserting the gene in reading frame to the secretory signal sequence MFα-factor. Removal of the putative SAS from the P45017α did, to a certain degree, prevent the integration of the modified protein into the *P. pastoris* membranes, as a stronger signal (68 kDa) in the concentrated

Stellenbosch University http://scholar.sun.ac.za

medium of the 96 hour fraction was observed. However the 55 kDa protein observed was not the P45017α protein since the same sized background protein was seen in the medium of the control strain.

The extracellular expression has therefore not proven to be very successful. Although the SAS of P45017α has been deleted, it is likely that other hydrophobic domains could still ensure binding to the yeasts membranes. Indeed, this had been seen with other truncated cytochromes P450 (76, 77). Pernecky and Coon have shown with P450 2E1 that there must be an additional membrane targeting domain beyond the first two hydrophobic regions, called S1 and S2, since the expressed protein was still found in the membrane fraction although the residues 3-48 were deleted (46). Clark and Waterman *et al*. have used the COS 1 heterologous expression system to model the targeting and insertion of bovine P45017α into the ER membrane *in vivo* and shown that deletion of the signal anchor domain did not abolish membrane association of the truncated P45017α, but resulted in a 5 fold reduction in insertion into the membrane (78).

Expression levels of cytochrome P45017α in KM71 (Mut^s) strains was not observed. Mut^s phenotype affects both the growth rate during induction and the accumulation of large amounts of endogenous alcohol oxidase. Mut^s strains may yield a higher proportion of correctly folded product in situations where folding is rate-limiting (79).

The secretion of human b_5 by *P. pastoris* was not observed. Neither in the medium nor in the cytoplasmic fraction of the yeast did we detect any histidine tagged b_5, which should have shown a protein band around 16 kDa. It has been observed that b_5 can aggregate into a dimer (32 kDa) or tetramere (64 kDa) (59), but no protein of identical molecular mass were detected. It could be possible that the histidine tag has either been digested by proteases or the access of the nickel-chelated probe to the histidine residues were obstructed by the folding of the protein. Polyclonal mouse anti-sheep b_5 antibodies were subsequently used but these antibodies also failed to detect any expressed b_5.

Western blot analysis of the cell culture medium of *P. pastoris* cells, expressing aromatase, indicated a protein with a molecular mass corresponding to 68 kDa after 72 and 96 hour induction phases. In the cytoplasmic fraction of these cells the aromatase

Stellenbosch University http://scholar.sun.ac.za

protein was seen in the 72 and 96 hour fractions, however, the molecular mass of the protein was 64 kDa. Although the expected molecular mass of aromatase would be around 58 kDa, the protein might be glycosylated or the α-factor secretion peptide was not successfully cleaved by proteolytic enzymes, which would explain the presence of a slightly larger molecular mass protein. The possibility that aromatase was glycosylated exist as aromatase has two potential glycosylation sites (Asn-12 and Asn-180) and Shimozawa *et al.* showed that the Asn-12 in the amino-terminal portion of aromatase is the site of glycosylation (54). Gartner *et al.* found a glycosylated aromatase expressed in baculoviral cells and has determined that the molecular weight of glycosylated aromatase was about 59 kDa compared to the 57.7 kDa unglycosylated aromatase (80). Interestingly, the first 20 amino acids of the aromatase sequence has been identified as an N-linked glycosylation consensus sequence with asparagine at position 12 of the glycosylation site while aromatase of human placenta has been shown to have attached sugar residues (81, 54, 82, 83). It is believed that the first 20 amino acid sequence with the sugar residues is protruding into the luminal space of the endoplasmic reticulum membrane with the hydrophobic region (residue 20 to 40) acting as the membrane spanning domain (80).

Several antibodies directed against aromatase have been prepared from purified human placental aromatase (20, 84). Polyclonal and monoclonal antibodies raised against a synthesised peptide, which is highly conserved within the sequence of aromatase, was also shown (85,86).

From the data presented here it seems that it would be very interesting to show activity of P45017α and aromatase in living cells. We have only screened the different clones for expressed proteins outside and inside the cells and have shown for P45017α and aromatase that they are produced and located intracellularly. By showing exactly where the proteins are located along the secretory pathway, i.e. ER, Golgi, secretory vesicle, periplasm or cell wall, we would be able to make better assumptions on the lack of cytochrome P450 export into the media. Tagging the expressed proteins with green fluorescent protein (GFP) could assist with the detection, although the additional size of the GFP protein could hamper the secretory process.

Stellenbosch University http://scholar.sun.ac.za

Based on this research, we can conclude that the full-length proteins, with their hydrophobic membrane domains intact, will not be readily secreted into the media. By deleting certain regions of the original cDNA, the protein can become more soluble and might then be secreted into the media. The process of optimising the extracellular secretion of P450-dependent enzymes will be extremely time consuming and it was decided to investigate the intracellular expression of the unmodified cytochromes P450 in *P. pastoris* instead.

Stellenbosch University http://scholar.sun.ac.za

REFERENCES

1 Swart, P., Swart, A. C., Waterman, M. R., Estabrook, R. W. and Mason, J. I. (1993) *J. Clin. Endocrinol. Metab.* **77**, 98-102.

2 Lin, D., Black, S. M., Nagahama, Y. and Miller, W. L. (1993) *Endocrinology* **132**, 2498-2506.

3 Lee-Robichaud, P., Wright, J.N., Akhtar, M. and Akhtar, M. (1995) *Biochem. J.* **308**, 901-908.

4 Auchus, R. J., Lee, T. C. and Miller, W. L. (1998) *J. Biol. Chem.* **273**, 3158-3165.

5 Thompson, Jr. E. A. and Siiteri, P. K. (1974) *J. Biol. Chem.* **249**, 5373-5378.

6 Miller, W. L., Auchus, J. and Geller, D. H. (1997) *Steroids* **62**, 133-142.

7 Auchus, R. J., Geller, D. H., Lee, T. C. and Miller, W. L. (1998) *Trends Endocrin. Met.* **9**, 47-50.

8 Onoda, M. and Hall, P. F. (1982) *Biochem. Biophys. Res. Commun.* **108**, 454-460.

9 Kominami, S., Ogawa, N., Morimune, R., Huang, D. Y. and Takemori, S. (1992) *J. Steroid Biochem.* **42**, 57-64.

10 Sasano, H. and Harada, N. (1998) *Endocr. Rev.* **19**, 593-607.

11 Chen, S. (1998) *Front. Biosci.* **3**, 922-933.

12 Chen, S., Zhou, D., Okubo, T., Kao, Y.-C. and Yang, C. (1999) *Endocr.-Relat. Cancer* **6**, 149-156.

13 Santen, R. J., Yue, W., Naftolin, F., Mor, G. and Berstein, L. (1999) *Endocr.-Relat. Cancer* **6**, 235-243.

14 Miller, W. R. (1999) *Endocr.-Relat. Cancer* **6**, 187-195.

15 Bulun, S. E., Zeitoun, K., Takayama, K., Noble, L., Michael, D., Simpson, E., Johns, A., Putman, M. and Sasano, H. (1999) *Endocr.-Relat. Cancer* **6**, 293-301.

16 Sasano, H., Sato, S., Ito, K., Yajima, A., Nakamura, J., Yoshihama, M., Ariga, K., Anderson, T. J. and Miller, W. R. (1999) *Endocr.-Relat. Cancer* **6**, 197-204.

17 Brodie, A. (2002) *Trends Endocrin. Met.* **13**, 61-65.

18 Lu, Q., Nakmura, J., Savinov, A., Yue, W., Weisz, J., Dabbs, D. J., Wolz, G. and Brodie, A. (1996) *Endocrinology* **137**, 3061-3068.

19 Osawa, Y. & Higashiyawa, T. (1980) *in "Microsomes, Drug Oxidations and Chemical Carcinogenesis"* (Coon, M. J., Conney, A. H., Estabrook, R. W., Gelboin, H. V., Gillette, J. R. & O'Brian, P. J., eds.), Vol. 1, pp. 225-228, Academic Press, New York.

20 Mendelson, C. R., Wright, E. E., Evans, C. T., Porter, J. C. & Simpson, E. R. (1985) *Arch. Biochem. Biophys.* **243**, 480-491.

21 Pasanen, M. and Pelkonen, O. (1981) *Biochem. Biophys. Res. Commun.* **103**, 1310-1317.

22 Tan, L. and Muto, N. (1986) *Eur. J. Biochem.* **156**, 243-250.

23 Nakajin, S., Shinoda, M. and Hall, P. F. (1986) *Biochem. Biophys. Res. Commun.* **134**, 704-710.

24 Hagerman, D. D. (1987) *J. Biol. Chem.* **262**, 2398-2400.

25 Kellis, J. T. and Vickery, L. E. (1987) *J. Biol. Chem.* **262**, 4413-4420.

26 Yoshida, N. and Osawa, Y. (1991) *Biochemistry* **30**, 3003-3010.

27 Imai, T., Globerman, H., Gertner, J., Kagawa, N. and Waterman M. R. (1993) *J. Biol. Chem.* **268**, 17317-17375.

28 Pompon, D., Liu, R. Y.-K., Besman M. J., Wang, P.-L., Shively, J. E. and Chen S. (1989) *Mol. Endo.* **3**. 1477-1487.

29 Swart, A. C., Swart, P., Roux, S. P.,van der Merwe, K. J., Pretorius, I. S. and Steyn, A. J. C. (1995) *Endocr. Res.* **21**, 289-295.

30 Zuber, M. X., Mason, J. I., Simpson, E. R. and Waterman, M. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 699-703.

31 Lähde, M., Raunio, H., Pelkonen, O., Karp, M. and Oker-Blom, C. (1993) *Biochem. Biophys. Res. Commun.* **197**, 1511-1517.

32 Sigle, R. O., Titus, M. A., Harada, N. and Nelson, S. D. (1994) *Biochem. Biophys. Res. Commun.* **201**, 694-700.

33 Gartner, C. A., Thompson, S. J., Rettie, A. E. and Nelson, S. D. (2001) *Protein Expres. Purif.* **22**, 443-454.

34 Romanos, M. A, Scorer, C. A and Clare, J. J. (1995) *DNA Cloning 2* (D. M. Glover and B. D. Hames, ed.) 2nd Ed, Oxford University Press Inc., New York.

35 Cregg, J. M. and Higgins, D. R. (1995) *Can. J. Bot.* **73**, S891-S897.

36 Hollenberg, C. P. and Gellisen, G. (1997) *Curr. Opin. Biotech.* **8**, 554-560.

37 Cregg, J. M., Vedvick, T. S. and Raschke, W. C. (1993) *Biotechnology (NY)* **11**, 905-910.

38 Gelissen, G. (2000) *Appl. Microbiol. Biot.* **54**, 741-750.

39 Lin Cereghino, J. L. and Cregg, J. M. (2000) *FEMS Microbiol. Rev.* **24**, 45-66.

40 Rossanese, O. W., Soderholm, J., Bevis, B. J., Sears, I. B., O'Connor, J., Williamson, E. K. and Glick, B. S. (1999) *J. Cell Biol.* **145**, 69-81.

41 Culvenor J. G., Henry, A., Artman, T., Evin, G., Galatis, D., Friedhuber, A., Rajiv Jayasena, U. L. H., Underwood, J. R., Beyreuther, K., Masters, C. L. and Cappai, R. (1998) *Int. J. Clin. Invest.* **5**, 79-89.

42 Lin Cereghino, G. P., Lin Cereghino, J., Ilgen, C. and Cregg, J. M. (2002) *Curr. Opin. Biotech.* **13**, 329-332.

Stellenbosch University http://scholar.sun.ac.za

43 Sakaguchi, M., Mihara, K. and Sato, R. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3361-3364.

44 von Wachenfeldt, C. and Johnson, E. F. (1995) *in "Cytochrome P450: Structure, Mechanism and Biochemistry"* (P. R. Ortiz de Montellano, ed.), 2nd ed., pp. 183-244. Plenum Press, New York.

45 Nelson, D. R. and Strobel, H. W. (1988) *J. Biol. Chem.* **263**, 6038-6050.

46 Pernecky, S. J. and Coon, M. J. (1996) *Method. Enzymol.* **272**, 25-34.

47 Li, Y. C. and Chiang, Y. L. (1991) *J. Biol. Chem.* **266**, 19186-19191.

48 Hsu, L.-C., Hu, M.-C., Cheng, H.-C., Lu, J.-C. and Chung, B.-C. (1993) *J. Biol. Chem.* **268**, 14682-14686.

49 Scheller, U., Kraft, R., Schröder, K.-L. and Schunck, W.-H. (1994) *J. Biol. Chem.* **269**, 12779-12783.

50 von Wachenfeldt, C., Richardson, T. H., Cosme, J. and Johnson, E. F. (1997) *Arch. Biochem. Biophys.* **339**, 107-114.

51 Krynetsky, E. Y., Drutsa, V. L., Kovaleva, I. E., Luzikov, V. N. and Uvarov, V. Y. (1993) *FEBS Lett.* **336**, 87-89.

52 De Nobel, J. G. and Barnett, J. A. (1991) *Yeast* **7**, 313-323.

53 Sudbery, P. E. (1996) *Curr. Opin. Biotech.* **7**, 517-524.

54 Shimozawa, O., Sakaguchi, M., Ogawa, H., Harada, N., Mihara, K. and Omura, T. (1993) *J. Biol. Chem.* **268**, 21399-21402.

55. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

56. Linder, S., Schliwa, M. and Kube-Granderath, E. (1996) *BioTechniques* **20**, 980-982.

57. Swart, P., Polson, A. and van der Merwe, K. J. (1985) *Prep. Biochem.* **15**, 1-8.

58. Laemmli, U. K. (1970) *Nature* **227**, 680-685.

59. Lombard, N., Swart, A. C., van der Merwe M. J. and Swart, P. (2002) *Endocr. Res.* **28**, 485-492.

60. Barnes, H. J., Arlotto, M. P. and Waterman, M. R. (1991*) Proc. Natl. Acad. Sci. USA* **88**, 5597-5601.

61. Larson, J. R., Coon, M. J. and Porter, T. D. (1991) *J. Biol. Chem.* **266**, 7321-7324.

62. Richardson, T. H, Hsu, M.-H., Kronbach, T., Barnes, H. J., Chan, G., Waterman, M. R., Kemper, B. and Johnson, E. F. (1993) *Arch. Biochem. Biophys*. **300**, 510-516.

63. Waterman, M. R. (1993) *Biochem. Soc. T.* **21**, 1081-1085.

64. Trant, J. M. (1996) *Gen. Comp. Endocr.* **102**, 173-82.

Stellenbosch University http://scholar.sun.ac.za

65. Pompon, D., Louerat, B., Bronine, A. and Urban, P. (1996) *Method. Enzymol.* **272**, 51-63.

66. Asseffa, A., Smith, S. J., Nagata, K., Gillette, J., Gelboin, H. V. and Gonzalez, F. J. (1989) *Arch. Biochem. Biophys.* **274**, 481-490.

67. Gonzalez, F. P., Aoyama, T. and Gelboin, H. V. (1991) *Methods Enzymol.* **206**, 85-92.

68. Zuber, M. X., Simpson, E. R. and Waterman, M. R. (1986) *Science* **234**, 1258-1261.

69. Corbin, C. J., Graham-Lorence, S., McPhaul, M., Mason, J. I., Mendelson, C. R. and Simpson, E. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8948-8952.

70. Trant, J. M., Lorence, M. C., Johnson, E. F., Shackelton, C. H. L., Mason, J. I. and Estabrook, R. W. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9756-9760.

71. Lorence, M. C., Trant, J. M., Clark, B. J., Khyatt, B., Mason, J. I., Estabrook, R. W. and Waterman, M. R. (1990) *Biochemistry* **29**, 9819-9824.

72. Crespi, C. L. (1991) *Method. Enzymol.* **206**, 123-129.

73. Trant, J. M. (1996) *Arch. Biochem. Biophys.* **326**, 8-14.

74. Boyle, S. M., Popp, M. P., Smith, W. C., Greenberg, R. M. and James, M. O. (1998) *Mar. Environ. Res.* **46**, 25-28.

75. Anderson, M. D., Busk, P. K., Svendsen, I and Møller, B. L. (2000) *J. Biol. Chem.* **275**, 1966-1975.

76. Pernecky, S. J., Larson, J. R., Philpot, R. M. and Coon, M. J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2651-2655.

77. Sagara, Y., Barnes, H. J. and Waterman, M. R. (1993) *Arch. Biochem. Biophys.* **304**, 272-278.

78. Clark, B. J. and Waterman, M. R. (1991) *J. Biol. Chem.* **266**, 5898-5904.

79. Cregg, J. M., Tschopp, J. K., Stillman, C., Siegel, R., Afong, M., Craig, W. S., Buckholz, R. G., Madden, K. R., Kellaris, P. A., Davis, G. R., Smiley, B. L., Cruze, J., Torregrossa, R., Veliçelebi, G. and Thill, G. P. (1987) *Biotechnology (NY)* **5**, 479-485.

80. Gartner, C. A., Thompson, S. J., Rettie, A. E. and Nelson, S. D. (2001) *Protein Expres. Purif.* **22**, 443-454.

81. Sethumadhavan, K., Bellino, F. L. and Thotakura, N. R. (1991) *Mol. Cell. Endocrinol.* **78**, 25-32.

82. Chen, S., Zhou, D., Swiderek, K. M., Kadohama, N., Osawa, Y. and Hall, P. F. (1993) *J. Steroid Biochem.* **44**, 347-350.

83. Amarneh, B., Corbin, C. J., Peterson, J. A., Simpson, E. R. and Graham-Lorence, S. (1993) *Mol. Endocrinol.* **7**, 1617-1624.

84. Washida, N., Kitawaki, J., Higashiyama, T., Matsui, S. (1996) *Steroids* **61**, 126-132.

Stellenbosch University http://scholar.sun.ac.za

85. Sanghera, M. K., Simpson, E. R., McPhaul, M. J., Kozlowski, G., Conley, A. J. and Lephart, E. D. (1991) *Endocrinology* **129**, 2834-2844.

86. Turner, K. J., Macpherson, S., Millar, M. R., McNeilly, A. S., Williams, K., Cranfield, M., Groome, N. P., Sharpe, R. M., Fraser, H. M. and Sauders, P. T. K. (2002) *J. Endocrinol.* **172**, 21-30.

Stellenbosch University http://scholar.sun.ac.za

PART IV

CLONING, INTRACELLULAR EXPRESSION AND PURIFICATION OF RECOMBINANT HUMAN CYTOCHROME P45017α (17α-HYDROXYLASE/17,20-LYASE) IN *PICHIA PASTORIS*

Stellenbosch University http://scholar.sun.ac.za

INTRODUCTION **202**

CHAPTER 1 : CONSTRUCTION OF THE FULL-LENGTH HUMAN CYP17 INTO PPIC3.5K EXPRESSION VECTOR AND INTEGRATION IN *P. PASTORIS* **203**

1.1 Materials and Methods **203**

1.1.1 Subcloning of the CYP17 cDNA 203

1.1.2 Construction of the pPIC3.5K-CYP17 expression vector 205

1.1.3 Screening of GS115/pPIC9K-CYP17 and KM71/pPIC9K-CYP17 transformants for integration 207

1.2 Results **207**

1.2.1 Subcloning of the CYP17 cDNA 207

1.2.2 Construction of the pPIC3.5K-CYP17 expression vector 208

1.2.3 Screening of GS115/pPIC3.5K-CYP17 and KM71/pPIC3.5K-CYP17 transformants for integration 209

CHAPTER 2 : CULTURE OF *P. PASTORIS* EXPRESSING P45017α AND SUCCESSFUL EXPRESSION OF THE PROTEIN **211**

2.1 Materials and Methods **211**

2.1.1 Expression of P45017α in *P. pastoris* by batch cultivation 211

2.1.2 Expression of P45017α in *P. pastoris* by large-scale batch cultivation 211

2.1.3 P45017α activity assay in yeast cells 212

2.1.4 Monitoring intracellular expression of P45017α 212

2.1.5 SDS-PAGE and Western blot analyses 213

2.1.6 Other Methods 213

2.2 Results **214**

2.2.1 Growth of *P. pastoris* strains 214

2.2.2 Substrate conversion of P45017α expressing cells 215

2.2.3 Expression of P45017α by *P. pastoris* GS115 216

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 3 : DEVELOPMENT AND EVALUATION OF METHODS FOR THE PREPARATION OF MICROSOMES FROM *P. PASTORIS* EXPRESSING HUMAN P45017α **217**

3.1 Materials and Methods **217**

3.1.1 Isolation of P45017α from microsomes 217

3.1.1.1 Preparation of yeast microsomes and isolation of P45017α with sonication in Emulgen 913 217

3.1.1.2 Preparation of microsomes in detergent and subsequent isolation of P45017α with Emulgen 913 or Triton X-100 218

3.1.2 Direct solubilisation of P45017α from *P. pastoris* 218

3.1.2.1 Lysis of *P. pastoris* and solubilisation of P45017α by in Emulgen 913 218

3.1.2.2 Lysis of *P. pastoris* in β-mercaptoethanol (β-ME) and solubilisation of P45017α by Emulgen 913 219

3.1.3 Reduced versus oxidised carbon monoxide induced difference spectra 219

3.1.4 Substrate-induced difference spectra 220

3.2 Results **220**

3.2.1 Preparation and solubilisation of P45017α from yeast microsomes 220

3.2.2 Direct solubilisation of P45017α from *P. pastoris* 223

3.2.3 Spectral properties of P45017α in solubilised and insoluble microsomes 224

CHAPTER 4 : THE DEVELOPMENT AND EVALUATION OF A PROTOCOL FOR THE PURIFICATION OF P45017α FROM MICROSOMES **228**

4.1 Materials and Methods **228**

4.1.1 Purification of Emulgen 913 solubilised P45017α 228

4.1.1.1 Ni^{2+}-NTA column chromatography 228

4.1.1.2 Size exclusion chromatography 228

4.1.2 Purification of Emulgen 913 solubilised P45017α and direct column loading 229

4.1.3 Purification of Emulgen 913/β-ME solubilised P45017α 229

4.1.4 Adsorption chromatography 230

Stellenbosch University http://scholar.sun.ac.za

4.1.5 Two-step Purification of Emulgen 913/β-ME solubilised P45017α 230

4.1.5.1 Detergent removal with Amberlite XAD resin followed by Ni^{2+}-NTA column chromatography 230

4.1.5.2 Ni^{2+}-NTA column chromatography followed by detergent removal with Amberlite XAD resin 231

4.2 Results 232

4.2.1 Purification of P45017α with Emulgen 913 and incubation overnight 232

4.2.2 Purification of P45017α with Emulgen 913 and direct column loading 233

4.2.3 Purification of Emulgen 913/β-ME solubilised P45017α 235

4.2.4 Adsorption chromatography 238

4.2.5 Two-step Purification of Emulgen 913/β-ME solubilised P45017α 238

4.2.5.1 Detergent removal followed by Ni^{2+}-NTA chromatography 238

4.2.5.2 Ni^{2+}-NTA chromatography followed by detergent removal 240

CHAPTER 5 : DISCUSSION 243

REFERENCES 244

Stellenbosch University http://scholar.sun.ac.za

INTRODUCTION

In the previous section, extracellular expression of a full-length human P45017α in *P. pastoris* produced insufficient amounts of P45017α proteins. Many factors, which could have contributed to the failure of secreting expressed proteins were previously discussed. There was, however, a possibility that P45017α would be expressed intracellulary in *P. pastoris*. Therefore, in this section we cloned and expressed the full-length human CYP17 cDNA into an intracellular, pPIC3.5K, expression vector.

We overexpressed of functional human P45017α in *P. pastoris*. The recombinant yeast strains were capable of converting progesterone to 17α-hydroxyprogesterone and 16α-hydroxyprogesterone. Different detergents and cell lysis methods were used to isolate the yeast microsomes and the P450 content in detergent-solubilised clear lysate and yeast microsomes was determined by reduced-carbon monoxide and substrate induced spectra. The intact hemeprotein with six histidine residues incorporated to the carboxyl terminus was purified by a two-step purification procedure to electrophoretic homogeneity.

This is the first expression of a human P450-dependent enzyme in *P. pastoris*. In addition, it is the first report of the overexpression of any P45017α with an unmodified amino terminal.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 1 : CONSTRUCTION OF THE FULL-LENGTH HUMAN CYP17 INTO pPIC3.5K EXPRESSION VECTOR AND INTEGRATION IN *P. PASTORIS*

1.1 Materials and Methods

1.1.1 Subcloning of the CYP17 cDNA

All DNA manipulations were carried out using standard procedures (1). The full-length human CYP17, located in the pPIC9K-CYP17 vector, was used as the DNA template to amplify a CYP17 as depicted in Figure 1.1. The upstream primer P105 (5'-ATGC<u>GAATTC</u>ATGTGGGAGCTCGTG-3') and the downstream primer P106 (5'-GTACGCGGCCGC<u>GAATTC</u>*TTA*GTGATGGTGAT-3') were used. The P105 and P106 primers each introduced an *Eco*RI site (underlined). P105 introduced a stop codon (italics) after the nucleotide bases encoding six histidine residues.

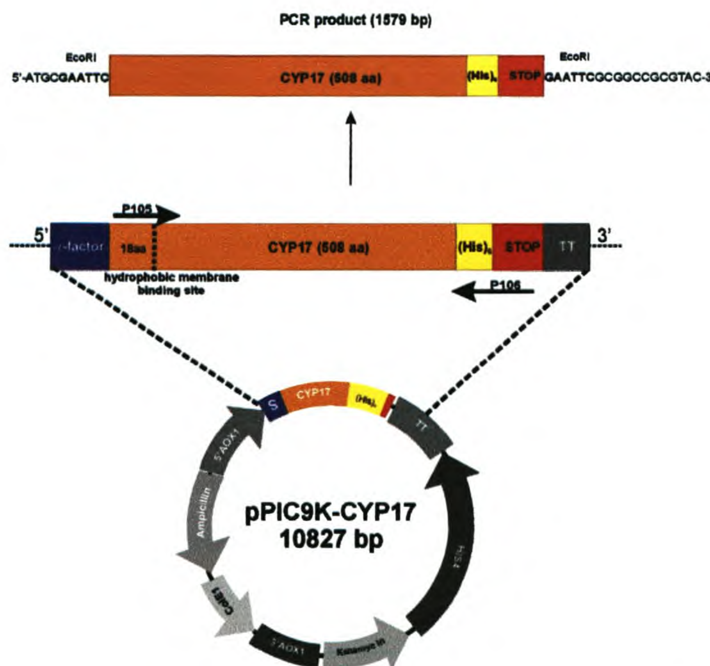

Figure 1.1 Schematic representation of the cloning strategy for the human CYP17 from the pPIC9K-CYP17 recombinant vector. The downstream primer P106 was homologous to the last four histidine sequences and the stop codon. Both primers had an *Eco*RI site for subcloning into the pUC18 vector.

Stellenbosch University http://scholar.sun.ac.za

Each PCR amplification mixture (100 μl) contained pPIC9K-CYP17 vector DNA (750 ng) and the reaction was carried out as described in Part III section 1.1.1.2. The amplified DNA fragment was electrophoresed on an 0.8% low melting agarose gel, extracted and purified as described in Part III section 1.1.1.2.

The CYP17 PCR product was digested with *Eco*RI and cloned into the pUC18 vector linearised with *Eco*RI, depicted in Figure 1.2. Two separate digestions, one containing amplified CYP17 DNA (2 μg) and the other containing pUC18 DNA (2 μg) were carried out as described in Part III section 1.1.1.2.2. The digested CYP17 fragment and linearised vector were electrophoresed on an 0.8% low-melting agarose gel in TAE buffer (40 mM) and purified as described in Part III section 1.1.1.2.2.

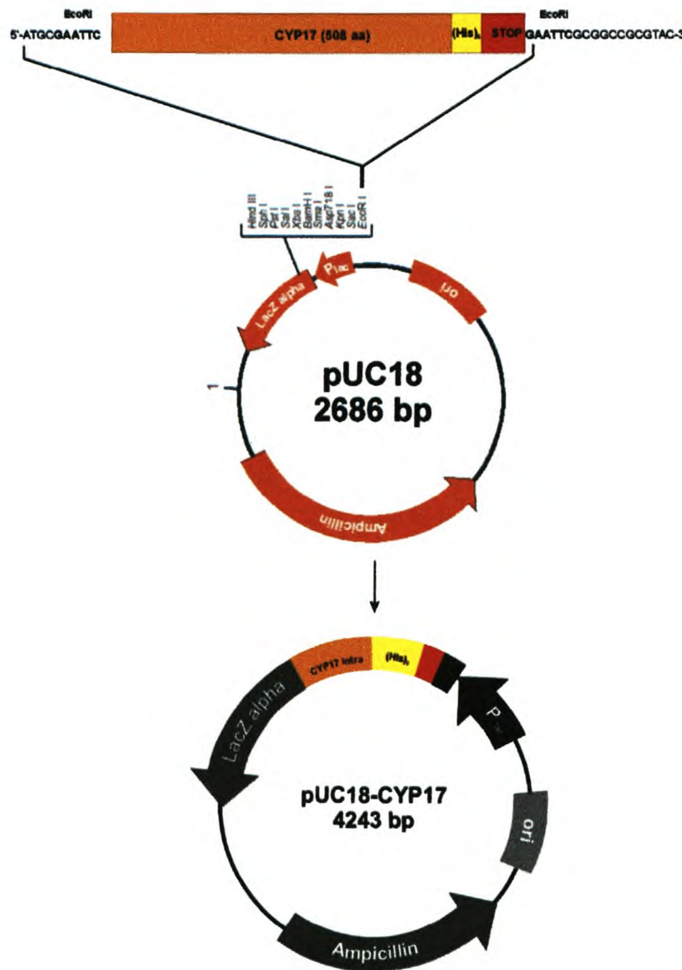

Figure 1.2 Schematic representation of the subcloning of the CYP17 into pUC18 vector.

Stellenbosch University http://scholar.sun.ac.za

The linearised pUC18 vector (3.66 μg) was dephosphorylated with CIAP(0.042 U) as described in Part III section 1.1.1.2.3. The purified digested CYP17 fragment (112 ng) and pUC18 (200 ng) DNA were subsequently ligated as described in Part III section 1.1.1.2.4. Competent JM109 cells were transformed with the ligated vectors. In addition to ampicillin selection, the vector contained the lacZ gene which allowed for blue/white colony selection.

Positive colonies were subsequently incubated overnight in LB medium (5 ml) at 37°C and the plasmid DNA was isolated using the Wizard™ *Plus* Minipreps DNA Purification System as described in Part III section 1.1.1.2.5. The purified recombinant plasmid DNA was digested with *Xba*I and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit as described in Part III section 1.1.1.2.5. The purified isolated plasmids were analysed by DNA sequencing and subsequently used in the construction of the pPIC3.5K-CYP17 expression vector.

1.1.2 Construction of the pPIC3.5K-CYP17 expression vector

The pUC18-CYP17 construct (4243 bp), was digested with *Eco*RI and the CYP17 was ligated into pPIC3.5K vector, linearised with *Eco*RI as depicted in Figure 1.3.

Restriction digestions of pUC18-CYP17 (2 μg) and pPIC3.5K vector DNA (2 μg) were carried at 37°C for 2 hours each in a final volume of 30 μl containing 5 μl restriction buffer B (supplied), acetylated BSA (5 μg) and *Eco*RI (10 U). The digested CYP17 fragment and linearised vector were electrophoresed on an 0.8% low-melting agarose gel, extracted and purified from the agarose as described in Part III section 1.1.1.2.2.

The linearised pPIC3.5K DNA (3.66 μg) was dephosphorylated with CIAP (0.012 U) and purified as described in Part III section 1.1.1.2.3. The purified digested CYP17 (35 ng) and pPIC9K DNA (200 ng) were subsequently ligated as described in Part III section 1.1.1.2.4. Competent JM109 cells were transformed with the ligated vectors and transformants were selected on LB-agar plates with ampicillin as described in Part III section 1.1.1.4.1. Positive colonies were subsequently incubated overnight in LB

Stellenbosch University http://scholar.sun.ac.za

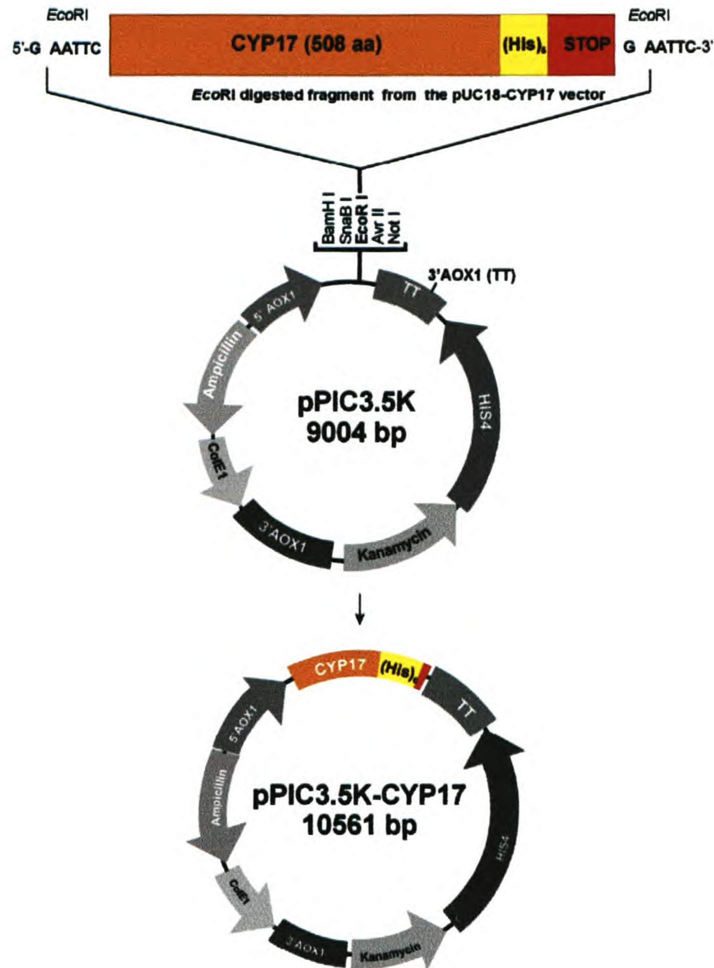

Figure 1.3 Schematic representation of the cloning strategy for the construction of the pPIC3.5K-CYP17 expression vector. The cloned CYP17 in the pUC18 vector was digested with *Eco*RI and inserted into the pPIC3.5K expression vector at the *Eco*RI *site*. The CYP17 was cloned in reading frame of the α-factor secretion signal and contains a nucleotide sequence encoding the six histidine tag.

medium (5 ml) at 37°C and the plasmid DNA was isolated as described in Part III section 1.1.1.2.5. The purified recombinant plasmid DNA was digested with *Xba*I and analysed by agarose gel electrophoresis. Plasmids containing the correct insert were subsequently prepared by a modified alkaline/SDS lysis method using a Nucleobond® AX100 plasmid isolation kit as described in Part III section 1.1.1.2.5.

The purified recombinant pPIC3.5K-CYP17 vector was confirmed by restriction digest analyses and DNA sequence analyses as described in Part III section 1.1.1.4.3. Primers 5'AOX1P and 3'AOX1P were used as illustrated in Figure 1.4.

Stellenbosch University http://scholar.sun.ac.za

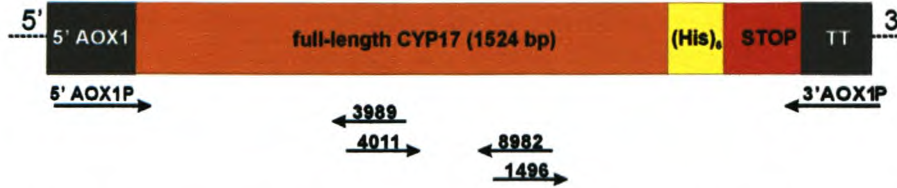

Figure 1.4 Schematic representation of the DNA sequencing strategy for the full-length human CYP17 in the pPIC3.5K-CYP17 expression vector.

1.1.3 Screening of GS115/pPIC9K-CYP17 and KM71/pPIC9K-CYP17 transformants for integration

The cloned expression vectors, pPIC3.5K-CYP17 (10 μg) and the parent vector pPIC3.5K (10 μg), were linearised with *Sal*I and transformed into competent *P. pastoris* strains, GS115 Mut^+ (Methanol utilisation plus) and KM71 Mut^s (Methanol utilisation slow) using the spheroplast method as described in Part III section 1.1.1.5. The presence of CYP17 cDNA integrated into the yeast chromosomes was subsequently confirmed by PCR amplification analyses as described in Part III section 1.1.1.5.

1.2 Results

1.2.1 Subcloning of the CYP17 cDNA

The full-length human CYP17, located in the pPIC9K-CYP17 expression vector was used as the template to amplify a CYP17, yielding a PCR product (1557 bp). This was confirmed by 1.0% agarose gel electrophoresis (Figure 1.5. This amplified product was subsequently subcloned into an intermediate vector, pUC18, before cloning it into the pPIC3.5K expression vector.

The integrity of the newly constructed pUC18-CYP17 vector was confirmed by restriction digest analyses with *Xba*I followed by 1% agarose gel electrophoresis. The pUC18-CYP17 plasmid (4243 bp) containing the insert yielded the expected two bands corresponding to 2970 and 1273 bp as shown in Figure 1.6, lanes 5 and 6.

Stellenbosch University http://scholar.sun.ac.za

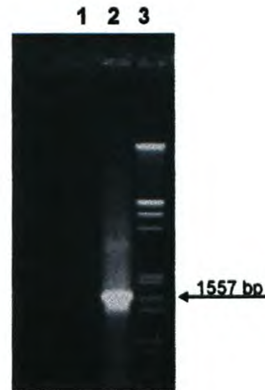

Figure 1.5 PCR amplification analysis of the amplified CYP17. The reaction products (10 µl) were analysed by 1% agarose gel electrophoresis. **Lane 1:** Negative control, PCR amplification without DNA template. **Lane 2:** Amplified CYP17 (PCR product, 1557 bp, using primers P105 and P106. **Lane 3:**DNA marker, λDNA/*Eco*RI/*Hind*III.

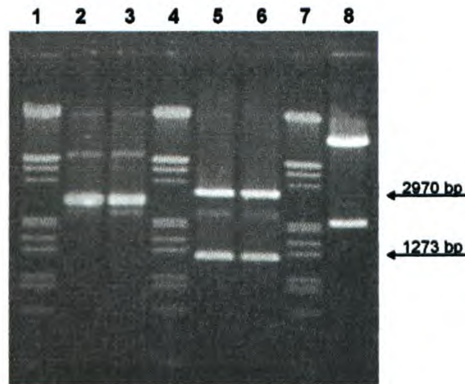

Figure 1.6 Restriction digest analysis of pUC18 plasmids containing the CYP17. Lanes 1, 4 and 7: DNA marker, λDNA/*Eco*RI/*Hind*III. **Lanes 2 and 3:** undigested plasmid DNA (0.5 µg). **Lanes 5 and 6:** *Xba*I digest of plasmid DNA (0.5 µg) containing the CYP17 yielding 2970 and 1273 bp fragments. **Lane 8:** Positive control containing *Xba*I digest of pPIC9K-CYP17 plasmid DNA (0.5 µg).

1.2.2 Construction of the pPIC3.5K-CYP17 expression vector

The *Eco*RI digested pUC18-CYP17 and pPIC3.5K plasmid DNA were electrophoresed on an 0.8% low melting agarose gel and the CYP17 and linearised plasmid DNA were excised and purified. Prior to ligation the cohesive compatible ends of the linearised pPIC3.5K plasmid were dephosphorylated and the CYP17 successfully ligated.

Transformed bacterial colonies were grown in small-scale cultures. Plasmid DNA isolated from each culture was digested with *Eco*RI and analysed by 1% agarose gel electrophoresis. The cloned pPIC3.5K-CYP17 vector (10561 bp) yielded the correct fragments of 8508 and 2053 bp as shown in Figure 1.7, lanes 13 and 15.

Stellenbosch University http://scholar.sun.ac.za

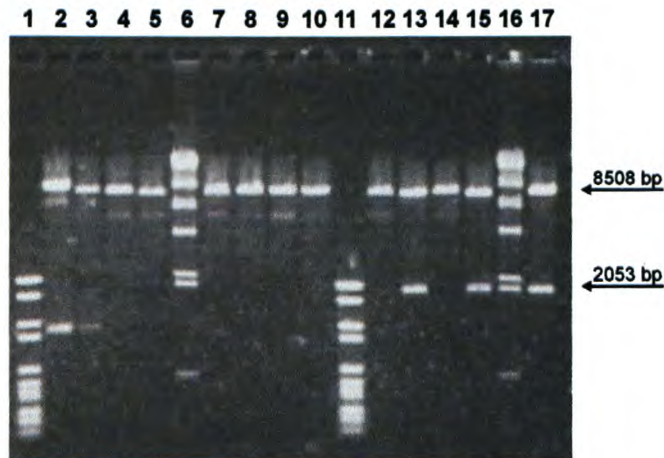

Figure 1.7 Restriction digest analysis of pPIC3.5K plasmids containing the CYP17. Lanes 1 and 11: DNA molecular weight marker VI (Roche). **Lanes 6 and 16:** DNA marker, λDNA/*Hind*III. **Lanes 2 and 3:** *Xba*I digest of plasmid DNA (0.5 μg) containing the CYP17 in the 3'-5' orientation. **Lanes 4-5, 7-10, 12 and 14:** *Xba*I digest of plasmid DNA (0.5 μg) without an insert. **Lanes 13 and 15:** *Xba*I digest of plasmid DNA (0.5μg) containing the CYP17 in the 5'-3' orientation yielding fragments 8508 and 2053 bp. **Lane 17:** Positive control containing *Sac*I digested pPIC9K plasmid DNA (0.5 μg).

DNA sequencing analyses, Appendix A.5, confirmed that the CYP17 had been inserted in reading frame for eukaryotic translation initiation and secretion.

1.2.3 Screening of GS115/pPIC3.5K-CYP17 and KM71/pPIC3.5K-CYP17 transformants for integration

PCR followed by 1% agarose gel electrophoretic analyses of lysed His^+ transformants showed that the CYP17 was integrated correctly, since two bands were produced. One corresponded to the CYP17 (1777 bp) and the other to *AOX1* gene (2100 bp) as shown in Figure 1.8, lane 3. A gene-specific primer STE1 together with 5'-AOX1P primer (Figure 1.8, lane 4) was also used to confirm the presence of CYP17. Mut^+ integrants with only the parent vector pPIC3.5K produced a 220 and 2100 bp PCR products as seen in Figure 1.8, lane 5.

The same PCR analyses were carried out on KM71 integrants. Two bands were produced. One corresponded to the CYP17 (1777 bp) and the other to the *AOX1* gene (4500 bp) as shown in Figure 1.9, lane 3.

Stellenbosch University http://scholar.sun.ac.za

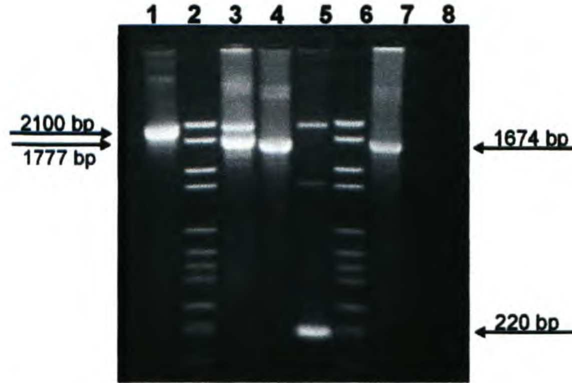

Figure 1.8 Direct PCR screening of GS115/pIC3.5K-CYP17 clones. Lane 1: Amplification of the wild-type *AOX1* gene (2100 bp) using AOX1P primers. **Lanes 2 and 6:** DNA molecular weight marker VI. **Lanes 3 and 4:** Contains a clone carrying both the *AOX1* gene (2100 bp) and CYP17 insert cloned into the pPIC3.5K plasmid resulting in a 1777 bp PCR product. In lane 3 AOX1P primers were used and in lane 4 insert-specific primer STE1 and the 3'AOX1P primer were used resulting in a 1674 bp PCR product. **Lane 5:** Contains a clone with the *AOX1* gene (2100 bp) carrying the pPIC3.5K vector without insert (220 bp). AOX1P primers were used. **Lane 7:** Positive PCR control containing CYP17 fragment (1674 bp) amplified from pPIC9K-CYP17 plasmid using AOX1P primers. **Lane 8:** Negative control, PCR amplification without DNA template. Each lane contained 20 µl reaction mixture.

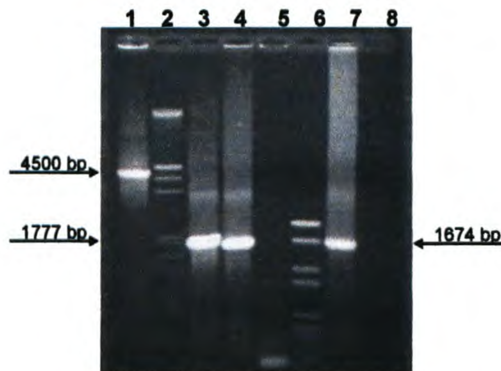

Figure 1.9 Direct PCR screening of KM71/pIC3.5K-CYP17 clones. The reaction products (20 µl) were analysed by 1% agarose gel electrophoresis. **Lane 1:** Amplification of the wild-type *AOX1* gene (4500 bp) using AOX1P primers. **Lane 2:** DNA marker, λDNA/*Eco*RI/*Hind*III. **Lanes 3 and 4:** Contains a clone carrying both the *AOX1* gene (4500 bp) and the CYP15 (1777 bp). In lane 3 AOX1P primers were used and in lane 4 a gene specific primer STE1 and the 3'AOX1P primer were used resulting in a 1674 bp fragment. **Lane 5:** Contains a clone carrying the parent pPIC3.5K vector without insert (220 bp). AOX1P primers were used. **Lane 6:** DNA molecular weight marker VI. **Lane 7:** Positive PCR control containing CYP17 fragment (1674 bp) amplified from pPIC9K-CYP17 plasmid using AOX1P primers. **Lane 8:** Negative control, PCR amplification without DNA template.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 2 : CULTURE OF *P. PASTORIS* EXPRESSING P45017α AND SUCCESSFUL EXPRESSION OF THE PROTEIN

2.1 Materials and Methods

2.1.1 Expression of P45017α in *P. pastoris* by batch cultivation

Clones that had been verified by direct PCR were tested for the expression of the desired protein. Single colonies of *P. pastoris* strains (GS115 and KM71) transformed with pPIC3.5K-CYP17 expression vectors were grown in BMGY (25 ml) to an optical density of 600 nm (OD_{600}) between 2 and 6 at 275 rpm for 20 hours at 30°C in an Innova 4000 shaking incubator (New Brunswick). Yeast cells were harvested by centrifugation at 2500 × g for 5 min at room temperature and resuspended to approximately OD_{600} = 1 in BMMY induction media (200 ml). Methanol-induced cultures were grown at 30°C with shaking at 275 rpm. Methanol (0.5% v/v final concentration) was added every 24 h and the induction was carried out up to 96 hours. To determine the optimal time for expression, aliquots (30 ml) of the culture media were removed at various time intervals i.e. 0, 24, 48, 72 and 96 hours. The expression culture was centrifuged at 2000 × g for 15 min at room temperature, the cell pellet was snap frozen in liquid nitrogen and stored at −80°C. These cells were tested for the presence of intracellularly expressed proteins by SDS-PAGE (Coomassie staining) and immunoblot analyses.

2.1.2 Expression of P45017α in *P. pastoris* by large-scale batch cultivation

GS115 cells from a freezer stock were inoculated in BMGY (25 ml), grown for 6 hours and added to BMGY media (200 ml), which was further incubated for 18 hours at 30°C at 250 rpm. Yeast cultures were centrifuged at 2000 × g for 5 min at room temperature and resuspended to an OD_{600} of between 0.6 and 1 in 4 times BMMY (250 ml). Methanol-induced cultures were grown at 30°C in a shaking incubator at 250 rpm. Methanol (0.5% (v/v) final concentration) was added to after every 24 h and the induction was carried out for up to 72 hours. The cells were harvested by centrifugation

Stellenbosch University http://scholar.sun.ac.za

at 2000 × g for 15 min at room temperature, weighed and resuspended in breaking buffer (BB). The cells were stored at –80°C until used for the preparation of microsomes.

2.1.3 P45017α activity assay in yeast cells

The metabolism of radiolabeled steroids was investigated in 48 h induced living yeast cells (OD_{600} =10, which is 5 × 10^7 cells/ml) in a shaking incubator at 30°C in a total volume of 2 ml. For each reaction, the cell suspension (4 ml) was harvested by centrifugation at 2000 × g for 5 min at room temperature, washed with phosphate buffered saline and resuspended in fresh BMMY induction media (1 ml). After a 15 min pre-incubation period at 30°C in an Erlenmeyer flask (25 ml), the reaction was initiated by adding an equal volume of BMMY medium supplemented with the appropriate tritium-labelled progesterone (600.000 cpm). The steroids present in the reaction media were extracted with methylene chloride (100:1 solvent to incubation mixture). The water phase was aspirated off and the methylene chloride was evaporated under a stream of nitrogen gas. The dried steroids were redissolved in methanol and analysed by HPLC (Waters, Milford,MA) equipped with in-line UV and radioactive (Radiomatic,Tampa, CA) detectors. Steroid metabolism was quantified by radioactivity.

2.1.4 Monitoring intracellular expression of P45017α

A rapid procedure was used to obtain a crude membrane fraction from cells grown by batch cultivation as described in section 2.1.1), to monitor specific P450 enzyme expression. Cell pellets, stored at –80°C, were thawed on ice and resuspended in BB (1 ml). All further steps were carried out at 4°C. The suspension was transferred to a 2 ml microcentrifuge tube and centrifuged at 3000 × g for 7 min at 4°C. The washed cells were resuspended in BB (0.5 ml) (OD_{600} of 100-125) and an equal amount of acid washed glass beads (Sigma G-9268, 425-600 μm) were added. The mixture was vortexed vigorously 10 times for 30 s each, with 30 s on ice between each vortexing. The supernatant was spun at 12000 × g for 10 min at 4°C to remove the glass beads and cell debris. The clear lysate (200-250 μl) was pipetted off carefully and stored at –20°C until analysed with SDS-PAGE and Western blot analyses.

Stellenbosch University http://scholar.sun.ac.za

2.1.5 SDS-PAGE and Western blot analyses

SDS-PAGE was carried out according to the method of Laemmli (2) using an acrylamide concentration of 10% (v/v). Low Range molecular weight markers (Biorad) and Rainbow coloured protein molecular weight markers (Amersham Pharmacia) consisting of myosin (220 kDa), phosphorylase b (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (21 kDa) and lysozyme (14.4 kDa). Wide range ColorBurst® Electrophoresis Markers (Sigma) were used alternatively to the Rainbow Markers, which have a 60 kDa protein marker instead of 66 kDa. Proteins were stained with Coomassie brilliant blue.

Electroblotting was carried out in a wet blotting system. Proteins were transferred onto a nitrocellulose membrane (Schleicher&Schuel) and the antigen-probe complex, which is a nickel (Ni^{2+}) activated horseradish-peroxidase (INDIA™ HisProbe-HRP, Pierce, Rockford, USA) was visualised by chemiluminescence (SuperSignal West HisProbe Kit, Pierce, Rockford. IL). The probe was diluted 1:5000 in TBST buffer. A six histidine tagged ureate oxidase (Pierce, Rockford, IL) was used as a positive control.

2.1.6 Other Methods

Protein concentrations were determined by the bicinchoninic acid method, BCA (Pierce Chemicals Co.) with bovine serum albumin as standard. Samples that contained β-mercaptoethanol interfered with the assay and had to be dialysed against MQ water to remove β-ME, before using the BCA assay.

Fractions collected after Ni^{2+}-NTA chromatography were concentrated and desalted with Ultrafree-15 Centrifugal filter devices (Millipore), according to the manufacturer's instructions. Samples with low concentrations of protein were concentrated with Ultrafree-CL centrifugal filter devices (Millipore) for optimal protein recovery.

Stellenbosch University http://scholar.sun.ac.za

2.2 Results

2.2.1 Growth of *P. pastoris* strains

GS115 transformants (Mut^+) were grown in BMGY media for 18 hours. They were harvested and resuspended in expression media (BMMY) at ~ OD_{600} of 1 and grown for four days supplementing with methanol every 24 hours. The cell density of cultures during the induction phase seems to increase linearly up to 48 hours after which it decreased into stationary phase as shown in Figure 2.1. The growth of the transformants was similar to the growth of the control strain GS115/β-Gal (Mut^+). The Control strain GS115/Albumin (Mut^s) had a lower initial growth rate, as expected, but after 48 hours the rate did not decrease as much as the wild-type methanol utilising strains (Mut^+) therefore reaching similar cell densities as the fast growing strains.

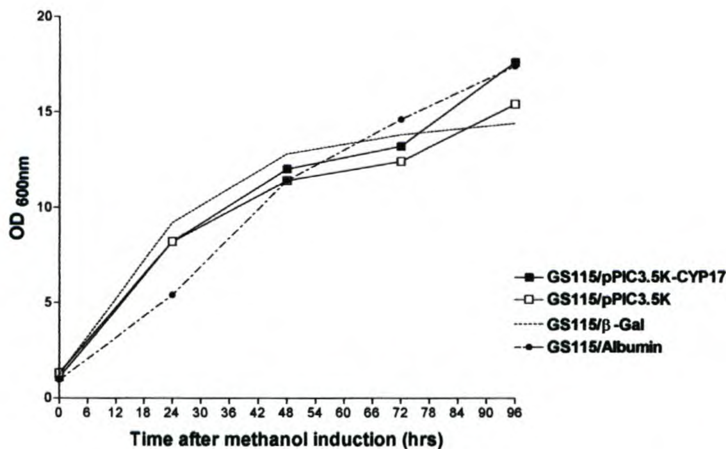

Figure 2.1 Growth curve of GS115 transformants during induction phase. Growth of the strains was monitored by optical density at 600 nm after the addition of methanol (0.5%, v/v final concentration). The medium was supplemented with methanol at 24 hour intervals. GS115/β-Gal is a Mut^+ control strain and WT is the wild-type GS115 strain.

KM71 transformants (Mut^S) were grown and induced the same way as the GS115 strains. The cultures grew in a linear fashion up to 48 hours, but then abruptly reached stationary phase as shown Figure 2.2.

It was therefore decided that the GS115 transformants would be used for further expression experiments.

Stellenbosch University http://scholar.sun.ac.za

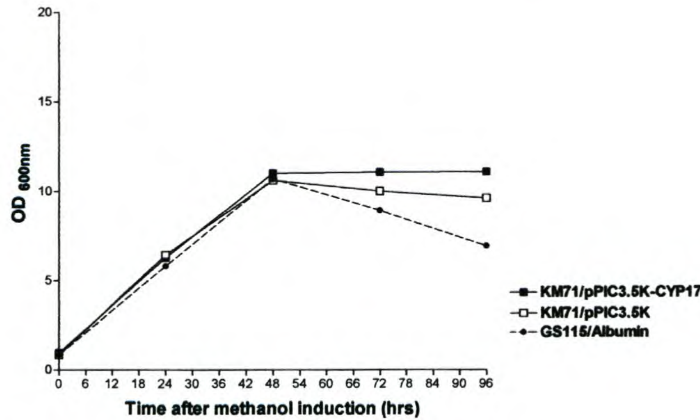

Figure 2.2 Growth curve of KM71 transformants during the induction phase. Growth of strains was monitored by OD_{600} after the addition of methanol (0.5%, v/v final concentration). The medium was supplemented with methanol at 24 hour intervals. GS115/Albumin is a Mut^s control strain.

2.2.2 Substrate conversion of P45017α expressing cells

P45017α was functionally expressed in *P. pastoris* as evidenced by the ability of recombinant yeast cells to convert progesterone to 17α-hydroxyprogesterone and 16α-hydroxyprogesterone as seen in Figure 2.3. At 4 min ≈ 55% of the progesterone had been metabolised, yielding ≈ 35% 17α-hydroxyprogesterone and ≈ 20% 16α-hydroxyprogesterone.

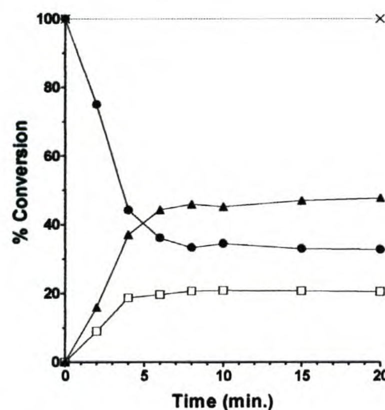

Figure 2.3 *In vivo* progesterone metabolism by yeast GS115 expressing human P45017α. Yeast cells, GS115, were transformed with pPIC3.5K expression plasmid containing the human CYP17 cDNA (pPIC3.5K-CYP17) or no cDNA insert (pPIC3.5K). 10 μM of ^{3}H-labeled progesterone was added to the cultures (1 × 10^9 cells/ ml). Aliquots of the culture were removed at certain time intervals, steroids extracted from the media and analysed by HPLC. ●, progesterone; ▲, 17α-hydroxyprogesterone; □, 16α-hydroxyprogesterone. ×, progesterone metabolised by GS115 cells transformed with pPIC3.5K expression plasmid containing no cDNA insert.

Stellenbosch University http://scholar.sun.ac.za

From 4 to 6 min the metabolism of progesterone gradually decreased and reached a stationary level of ≈ 65% progesterone metabolised. After 6 min the hydroxylated steroids also reached a maximum yield of ≈ 45% for 17α-hydroxyprogesterone and ≈ 20% for 16α-hydroxyprogesterone. The metabolites were measured for up to 2 hours and remained at the same levels. The assay was repeated several times resulting in similar substrate conversion rates. No conversion of progesterone was detectable using *P. pastoris* cells transformed with the pPIC3.5K vector containing no insert.

2.2.3 Expression of P45017α by *P. pastoris* GS115

The effect of duration of methanol induction (1-4 days) on the expression of the human P45017α is illustrated in Figure 2.4. Western blot analysis of the rapid membrane preparation of GS115 transformants, grown by batch cultivation (described in section 2.1.1), indicate the expression of P45017α from 24 to 96 h. Each lane contained the lysate of an equal number of cells (OD_{600} =1= 5 × 10^7) indicating a slight increase in P45017α expression per day was observed, although a protein band at a molecular mass corresponding to 40 kDa appears at 72 h and increases over time. This might be the result of the degradation of P45017α.

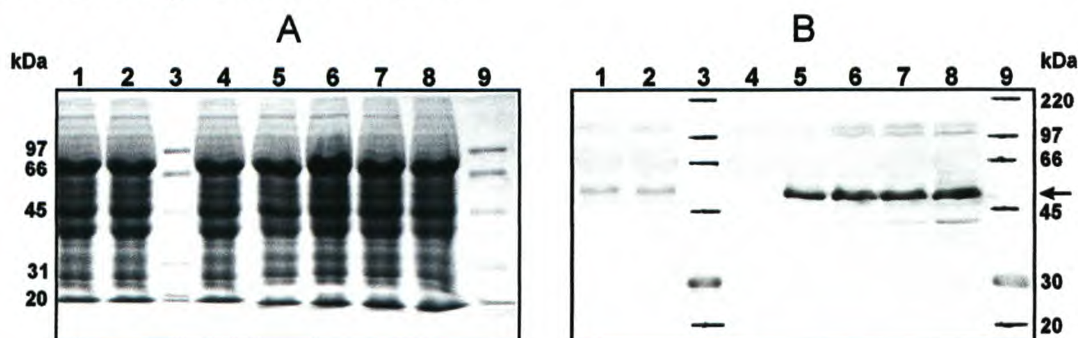

Figure 2.4 SDS-PAGE (A) and Western blot (B) analyses of total intracellular yeast proteins of GS115 cells induced to express P45017α for four days . Lanes 1 and 2: GS115 cells transformed with pPIC3.5K vector at 0 and 96 h, respectively. **Lanes 3 and 9:** Low-molecular weight Marker. **Lanes 4-8:** GS115 cells transformed with pPIC3.5-CYP17 vector at 0, 24, 48, 72 and 96 hours. Each lane contained 20 μg total proteins. P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 3 : DEVELOPMENT AND EVALUATION OF METHODS FOR THE PREPARATION OF MICROSOMES FROM *P. PASTORIS* EXPRESSING HUMAN P45017α

3.1 Materials and Methods

Two different methods (section 3.1.1 and section 3.1.2) were used to isolate P45017α from *P. pastoris*. The first method was to prepare microsomes (with or without Emulgen 913) from *P. pastoris* and to extract P45017α from the prepared microsomes with or without sonication in the presence of different detergents. The second method was to omit the preparation of the microsomes, but to directly extract the P45017α from the crude intracellular extract.

3.1.1 Isolation of P45017α from microsomes

3.1.1.1 Preparation of yeast microsomes and isolation of P45017α with sonication in Emulgen 913

All steps were carried out at 4°C. A 72 h cell pellet (5 ml, 2 g wet weight), stored at –20°C, was thawed on ice and resuspended in buffer A (2 ml)(50 mM KP_i, pH 7.6, 20% glycerol and 1 mM PMSF). An equal volume of acid washed glass beads (Sigma G-9268, 425-600 μm) was added. The cells were broken by vortexing (12 × 30 s) at 4°C with intermediate cooling on ice. The lysate was centrifuged at 12.000 × g for 10 min at 4°C to remove the cell debris and glass beads and the resulting supernatant (crude cell extract) was centrifuged at 165.000 × g for 1 h at 4°C to recover a microsomal pellet. The soluble fraction (crude lysate), was separated from the pellet and kept for SDS-PAGE analysis. The insoluble microsomal pellet was resuspended in buffer A (2 ml) with Emulgen 913 (1% v/v) and homogenised with a teflon rod. The lysate was sonicated (8 × 30 s on with 30s intervals) on ice and centrifuged at 200.000 × g for 1 h at 4°C to sediment the microsomes. The solubilised microsomal proteins and insoluble microsomal membranes, which were resuspended in buffer A with Emulgen 913 (1%

Stellenbosch University http://scholar.sun.ac.za

v/v), were stored at –80°C and used for SDS-PAGE and Western blot analyses. The same protocol was used for the preparation of microsomes from the wild type and control strains.

3.1.1.2 Preparation of microsomes in detergent and subsequent isolation of P45017α with Emulgen 913 or Triton X-100

All steps were carried out at 4°C. A 72 h cell pellet (5 ml, 2 g wet weight), stored at –20°C was thawed on ice and resuspended in buffer A (5 ml). The washed cells were harvested at 3000 × g for 10 min and resuspended in cell lysis buffer (50 mM sodium phosphate, pH 7.6, 20% glycerol, 1% v/v Emulgen 913 and 1 mM PMSF). An equal volume of acid washed glass beads (Sigma G-9268, 425-600 μm) was added. The cells were broken by vortexing (8 × 1 min with 1 min intervals) and sonication (8 × 30 s with 30 s intervals). The lysate was centrifuged at 12000 × g for 15 min at 4°C to remove the cell debris and glass beads and the resulting supernatant (crude cell extract) was centrifuged at 165.000 × g for 1 h at 4°C to recover a microsomal pellet. The soluble fraction (crude lysate), was separated from the pellet and kept for SDS-PAGE analysis.

The insoluble microsomal pellet was resuspended in buffer A, homogenised in a Teflon homogeniser and divided into two equal volumes. Emulgen 913 (1% v/v) was added to one microsomal aliquot and Triton-X100 (1% v/v) was added to the other. The two suspensions were stirred at 4°C for 2 hours and then centrifuged at 165000 × g for 1 hour at 4°C. The solubilised microsomal proteins were stored at –80°C and used for SDS-PAGE and Western blot analyses.

3.1.2 Direct solubilisation of P45017α from *P. pastoris*

3.1.2.1 Lysis of *P. pastoris* and solubilisation of P45017α by in Emulgen 913

All steps were carried out at 4°C. A 72 h cell pellet (20 ml, 0.5 g wet weight/ml), stored at –80°C was thawed on ice and resuspended in buffer A (10 ml). The cells were harvested by centrifugation at 2500 × g for 10 min at 4°C and resuspended in cell lysis buffer (8 ml). An equal volume of acid washed glass beads (Sigma G-9268, 425-600

Stellenbosch University http://scholar.sun.ac.za

μm) was added. The cells were broken by vortexing (8 × 1 min with 3 min intervals). The mixture was stirred for 20 min to ensure proper breakage of cells. The lysate was centrifuged at 12.000 × g for 15 min at 4°C to remove the cell debris and glass beads and the resulting supernatant (crude cell extract) was centrifuged at 165000 × g for 1 h at 4°C to separate solubilised proteins and insoluble microsomal membranes. These preparations were stored at −80°C and used for SDS-PAGE and Western blot analyses.

3.1.2.2 Lysis of *P. pastoris* in β-mercaptoethanol (β-ME) and solubilisation of P45017α by Emulgen 913

All steps were carried out at 4°C. A 72 h cell pellet (10 ml, 0.5 g wet weight/ml), previously stored at −80°C, was thawed on ice and resuspended in buffer A (1 ml per 0.5 g cells wet weight). The washed cells were harvested by centrifugation at 2000 × g for 10 min at 4°C and resuspended in cell lysis buffer (1 ml per 0.5 g cells wet weight) containing 1 mM EDTA and 10 mM β-ME. An equal volume of acid washed glass beads (Sigma G-9268, 425-600 μm) was added. The cells were broken by vortexing eight times for 1 min, with chilling on ice for 1 min between vortexing. The cell debris, nuclei, mitochondria and glass beads were removed by centrifugation at 14000 × g for 30 min at 4°C. Emulgen 913 (1% v/v) was added to the crude cell extract and the mixture was stirred for 4 hours at 4°C to ensure proper solubilisation of P45017α. The crude cell extract was centrifuged at 165000 × g for 1 h at 4°C to separate the solubilised P45017α from the insoluble microsomal pellet. These preparations were stored at −80°C and used for SDS-PAGE, Western blot and carbon monoxide (CO) difference spectra analyses.

3.1.3 Reduced versus oxidised carbon monoxide induced difference spectra

The cytochrome P450 concentration in the solubilised lysate and insoluble microsomal pellet was determined spectrophotometrically as described by Omura and Sato (3). The sample (1 ml) was diluted in phosphate buffer (1 ml)(0.1 M, pH 7.0) and the mixture was saturated with carbon monoxide by bubbling for 30 s on ice. The homogenate was divided into two equal volumes in two optically matched cuvettes. After an incubation

Stellenbosch University http://scholar.sun.ac.za

period of 3 min at 25°C a baseline was recorded between 500 and 400 nm in a dual-beam Cary 100 spectrophotometer (Varian, USA). Sodium dithionite was added to the sample cuvette and the contents mixed until the reduction of the cytochrome P450 was complete. A difference spectrum was recorded between 500 and 400 nm and after baseline correction the concentration of cytochrome P450 in the sample calculated using the extinction coefficient of 91 $mM^{-1}cm^{-1}$ for the difference in absorbance between 490 and 450 nm (3).

3.1.4 Substrate-induced difference spectra

Solubilised microsomal lysate (1 ml, 4.54 mg/ml), with a P450 enzyme concentration of 0.95 μM, was diluted in phosphate buffer (1 ml) (0.1 M, pH 7.0). The suspension was incubated for 5 min at 25°C, equally divided into two optically matched cuvettes and a baseline was recorded between 450 and 350 nm. Progesterone or pregnenolone, dissolved in ethanol, was added to the sample cuvette and an equal volume of ethanol to the reference cuvette. The final concentration of both steroids was 10 μM and the total volume of the additions did not exceed 0.2% of the volume in the cuvette. The contents of both cuvettes were mixed and a spectrum was recorded between 500 and 350 nm

3.2 Results

3.2.1 Preparation and solubilisation of P45017α from yeast microsomes

The first attempt to isolate P45017α from the yeast cells was to use mechanical breakage of cells with glass beads (see Materials and Methods 3.1.1.1). The cell lysate was subjected to differential centrifugation to separate the crude extract, containing soluble proteins (supernatant) from the insoluble microsomes (pellet). A small amount P45017α was solubilised in this step (Figure 3.1, lane 10), but the majority of P45017α was in the microsomal fraction, indicating microsomal localisation of P45017α (Figure 3.1, lane 8). Recombinant P45017α was isolated from *P. pastoris* microsomes by initial solubilisation in Emulgen 913 (1% v/v) followed by sonication and ultracentrifugation to recover solubilised P45017α (Figure 3.1, lane 7) from the insoluble microsomal

Stellenbosch University http://scholar.sun.ac.za

membrane (Figure 3.1, lane 8). No P45017α could be solubilised from the microsomal membrane in this manner. The yield of microsomes was 2-3 mg per 1 g of wet weight cells, which was too low for microsomal extractions from *P. pastoris*.

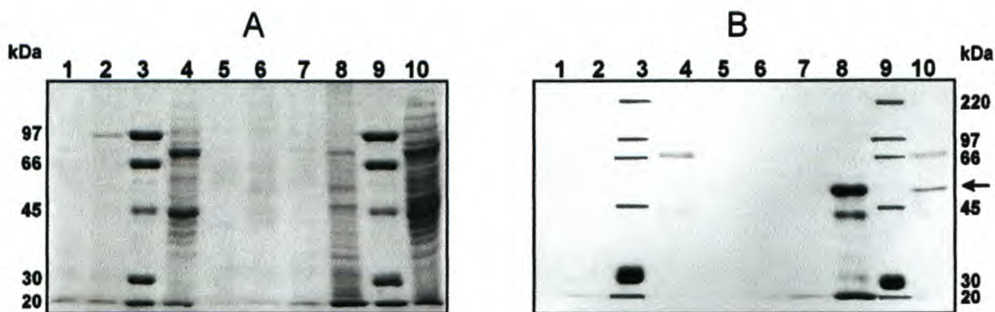

Figure 3.1 SDS-PAGE (A) and Western blot (B) analyses of proteins isolated from *P. pastoris* GS115/pPIC3.5K-CYP17 microsomes with Emulgen 913 and sonication. Lanes 1 and 2: solubilised microsomal proteins and insoluble microsomal membranes from GS115/3.5K cells, respectively. **Lanes 3 and 9:** Rainbow Marker. **Lane 4:** crude lysate (20 μg) from GS115/3.5K cells. **Lanes 5 and 6:** solubilised microsomal proteins and insoluble microsomal membranes from wild-type GS115 cells, respectively. **Lanes 7 and 8:** solubilised microsomal proteins (8 μg) and insoluble microsomal membranes (8 μg) from GS115/3.5K-CYP17 cells, respectively. **Lane 10:** crude lysate (20 μg) from GS115/3.5K-CYP17 cells. P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

The yeast cells were subsequently lysed by vortexing with glass beads in the presence of Emulgen 913 (1% v/v) (see Materials and Methods 3.1.1.2) followed by sonication. The idea was that vortexing would mechanically disrupt the cell wall of the yeast and the detergent would stabilise P45017α in the crude lysate and prevent the hydrophobic protein to interact with the microsomal membrane. Recombinant P45017α was isolated subsequently from *P. pastoris* microsomes by initial solubilisation in Emulgen 913 (1% v/v) and ultracentrifugation to recover solubilised P45017α (Figure 3.2, lane 6) from the insoluble microsomal membrane (Figure 3.2,lane 7). Soluble and insoluble microsomal fractions were also isolated from slow growing yeast KM71/3.5K-CYP17 (Figure 3.2 lanes 4 and 5) and the GS115/9K-CYP17 (Figure 3.2 lanes 9 and 10), which is an extracellular expression construct. Equal amounts of cells (wet weight) were used for the isolations. About 10-20% of P45017α (58 kDa) could be solubilised, but most of the protein stayed in the microsomal membrane. The slow growing yeast strain KM71 had no solubilised P45017α (Figure 3.2, lane 4 A and B). Comparing the GS115 to the KM71 strain, more protein was expressed by the fast methanol utilising strain. Interestingly, the extracellular expression clone GS115/9K-CYP17 contained the

Stellenbosch University http://scholar.sun.ac.za

expected protein band (P45017α) at a molecular mass corresponding to 58 kDa (lane 10, A and B), but an additional protein at a molecular mass corresponding to 60 kDa in the microsomal membrane (Figure 3.2, lane 10 A and B) was present. This could be a glycosylated form of P45017α, since the protein is translated under the secretory promoter (MFα-1).

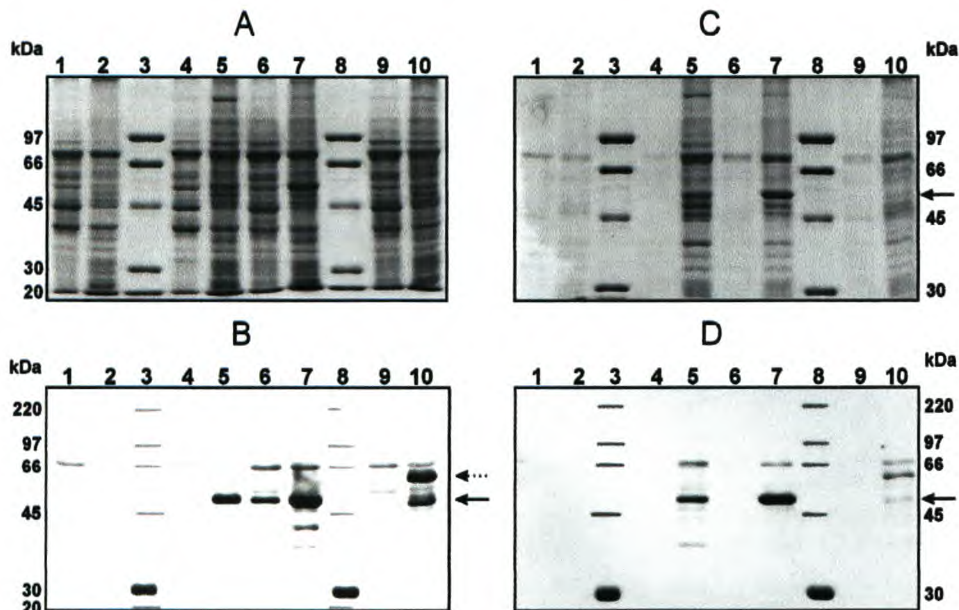

Figure 3.2 SDS-PAGE (A and C) and Western blot (B and D) analyses of proteins isolated from *P. pastoris* GS115/pPIC3.5K-CYP17 and KM71/pPIC3.5K-CYP17 microsomes with Emulgen 913 and sonication. (A) represents the first round of isolation of P45017α and (C) represents the re-isolation of P45017α from previously isolated insoluble microsomes: Lanes 1 and 2: solubilised microsomal proteins and insoluble microsomal membranes from GS115/3.5K cells, respectively. **Lanes 3 and 8:** Molecular Weight Marker. **Lanes 4 and 5:** solubilised microsomal proteins and insoluble microsomal membranes from KM71/3.5K-CYP17 cells, respectively. **Lanes 6 and 7:** solubilised microsomal proteins and insoluble microsomal membranes from GS115/3.5K-CYP17 cells, respectively. **Lanes 9 and 10:** solubilised microsomal proteins and insoluble microsomal membranes from GS115/9K-CYP17 cells, respectively. Each lane contained 20 and 5 μg of total protein in the left and right SDS-PAGE, respectively. P45017α (58 kDa) and potentially glycosylated P45017α are indicated by the solid and dotted arrows, respectively. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

The solubilisation method to yield P45017α from the yeasts microsomes was very poor, therefore an optimisation of the method was investigated (see Materials and Methods 3.1.1.2). Triton X-100 was used as an alternative to Emulgen 913. The comparison showed that neither detergents increased the yield of soluble P45017α from microsomes significantly (see Figure 3.3).

Stellenbosch University http://scholar.sun.ac.za

3.2.2 Direct solubilisation of P45017α from *P. pastoris*

As sonication might produce insoluble miscellular vesicles, impossible to solubilise, it was decided to add the detergent to the lysis buffer together with glass beads (see Materials and Methods 3.1.2) and instead of sonication a gentle stirring of the crude extract was attempted. Results show that approximately 60% of P45017α was in the solubilised supernatant (Figure 3.4; lane 6). 176 mg of total soluble protein (8.82 mg/ml)

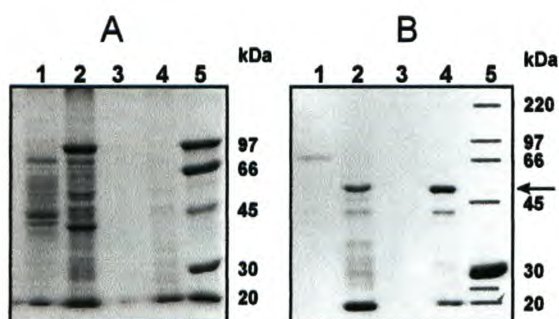

Figure 3.3 SDS-PAGE (A) and Western blot (B) analyses of P45017α solubilied from *P. pastoris* microsomes with different detergents. Lanes 1 and 2: Emulgen 913 solubilised microsomal proteins and insoluble microsomal membranes from GS115/3.5K-CYP17 cells, respectively. **Lanes 3 and 4:** Triton X-100 solubilised microsomal proteins and insoluble microsomal membranes from GS115/3.5K-CYP17 cells, respectively. Each lane contained 10 μg total protein.. **Lane 5:** Molecular Weight Marker. P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

and 16.92 mg of total insoluble microsome (8.46 mg/ml) was isolated, which produced 2.85 mg of soluble proteins per 1 g cells wet weight.

The final change in protocol, to optimise the isolation and solubilisation of P45017α from yeast cells, was to add β-ME to the cell lysis buffer (see Materials and Methods 3.1.2.2). Emulgen 913 (1% v/v) was added to crude extract and the lysate stirred for 4 hours. This gentle agitation (instead of sonication) was found to recover 70% of the P45017α as solubilised protein (Figure 3.5, lane 5), but a significant amount of P45017α was still in the insoluble microsomal membrane (Figure 3.5, lane 6). The yield of total soluble protein (4.54 mg/ml) was 227 mg while 33.5 mg of total insoluble microsomes (6.70 mg/ml) was isolated. The yield of microsomes was 2-3 mg per 1 g of wet wt cells.

Stellenbosch University http://scholar.sun.ac.za

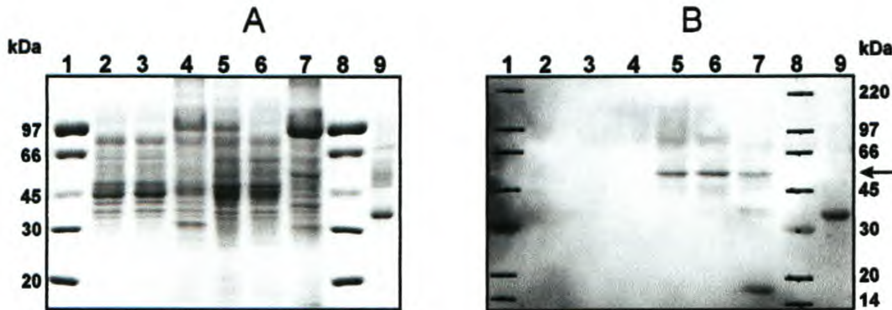

Figure 3.4 SDS-PAGE (A) and Western blot (B) analyses of proteins solubilised by Emulgen 913 from *P. pastoris* cells. Lanes 1 and 8: Molecular weight markers. **Lane 2:** crude cell extract from GS115/3.5K cells. **Lanes 3 and 4:** solubilised proteins and insoluble microsomal membranes from GS115/3.5K cells, respectively. **Lane 5:** crude cell extract from GS115/3.5K-CYP17 cells. **Lanes 6 and 7:** solubilised proteins and insoluble microsomal membranes from GS115/3.5K-CYP17 cells, respectively. **Lane 9:** Positive control containing a histidine tagged ureate oxidase (10 μg). Each lane contained 20 μg total protein. P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

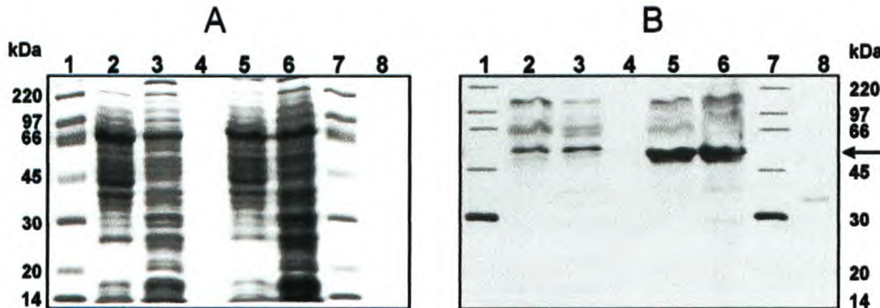

Figure 3.5 SDS-PAGE (A) and Western blot (B) analyses of proteins solubilised by β-mercaptoethanol and Emulgen 913 from *P. pastoris* cells. Lanes 1 and 7: Low molecular weight markers. **Lanes 2 and 3:** solubilised proteins and insoluble microsomal membranes from GS115/3.5K cells, respectively. **Lanes 5 and 6:** solubilised proteins and insoluble microsomal membranes from GS115/3.5K-CYP17 cells, respectively. **Lane 8:** Positive control containing a histidine tagged ureate oxidase (10 μg). Each lane contained 20 μg total protein. P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

3.2.3 Spectral properties of P45017α in solubilised and insoluble microsomes

A typical cytochrome P450 CO-induced difference spectrum with an absorption maximum of 448 nm was recorded for the soluble and insoluble preparation of microsomes from P45017α expressing GS115 cells as shown in Figure 3.6. No P450 enzyme was detected in any of the fractions from control cells (pPIC3.5K vector without cDNA). Figure 3.6A shows the reduced CO-induced difference spectra of detergent-solubilised proteins with β-ME. Only one peak at 448 nm, indicating the presence of a

Stellenbosch University http://scholar.sun.ac.za

hemeprotein P45017α, was recorded with a small shoulder at 427 nm. The heme group was not lost in the presence of β-mercaptoethanol (10 mM), which can occur occasionaly (4, 5). The yield of P450 enzyme in the solubilised cell lysate was approximately 50 nmol/L of cell culture at a P450 enzyme concentration of 0.95 μM. Figure 3.6B shows the reduced CO-induced induced difference spectra of solubilised proteins without Emulgen 913 and β-ME. An additional peak at 427 nm, indicating cytochrome oxidase, was recorded. The P450 enzyme was in a stable form since no absorption maximum around 420 nm was observed although it might be slightly obscured by the native mitochondrial cytochrome oxidase (425-427 nm).

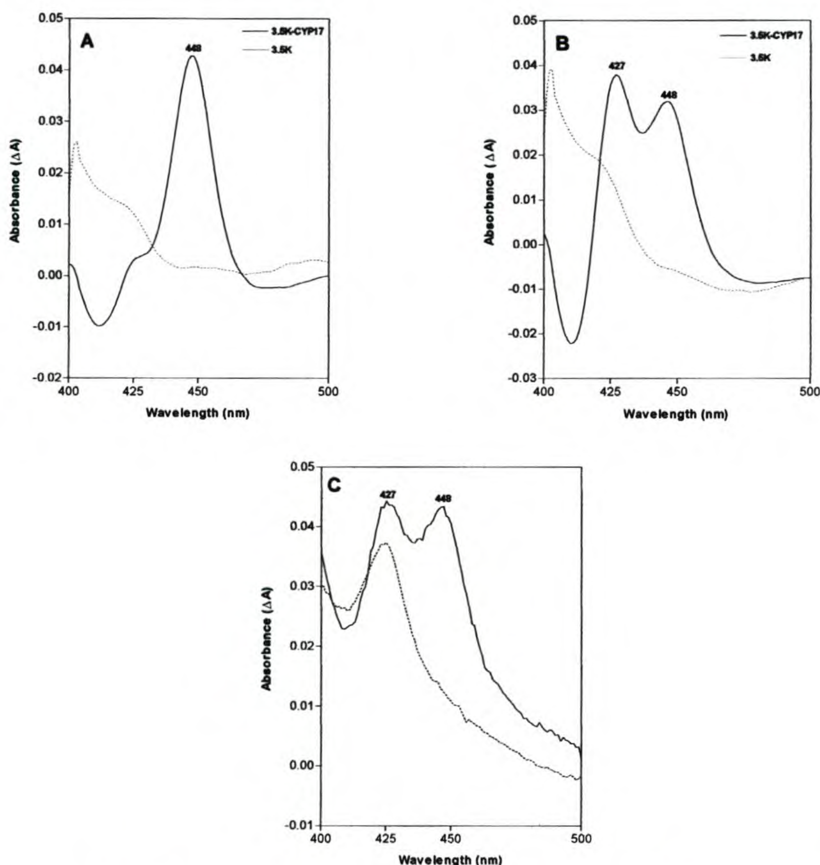

Figure 3.6 Reduced CO difference spectra of solubilised proteins (A and B) and microsomal fraction (C) of *P. pastoris* transformed with either the parent pPIC3.5K plasmid (·······) or recombinant pPIC3.5K-CYP17 vector (——). (A) Soluble protein fraction isolated in buffer A (100 mM sodium phosphate, pH 7.4, 20% v/v glycerol, 1 mM EDTA,1 mM PMSF, 10 mM β-mercaptoethanol and 1% v/v Emulgen 913). **(B)** Soluble protein fraction isolated in buffer BB (50 mM NaH_2PO_4, pH 7.4, 5% v/v glycerol, 1 mM EDTA 1 mM PMSF and 1% v/v Emulgen 913). **(C)** Insoluble microsomal fraction isolated in buffer BB (50 mM NaH_2PO_4, pH 7.4, 5% v/v glycerol, 1 mM EDTA, 1 mM PMSF and 1% v/v Emulgen 913). The samples were diluted 1:1 in 100 mM sodium phosphate buffer, pH 7.0, which was saturated with carbon monoxide. 2 mg/ml of solubilised protein was used.

Stellenbosch University http://scholar.sun.ac.za

This interference had previously been reported (29,6). To minimise the interference of mitochondrial cytochrome oxidase (peak at 427 nm), Emulgen 913 with β-ME is added to the lysis buffer. This addition will lower the 427 nm peak, as seen in Figure 3.6A, therefore exposing any P420 if present. Figure 3.6C shows the reduced CO-induced difference spectra of microsomal fraction resuspended in buffer BB and the P450 enzyme concentration was calculated to be 0.70 μM. The degradation of detergent-solubilised P45017α after 2 months is shown in Figure 3.7.

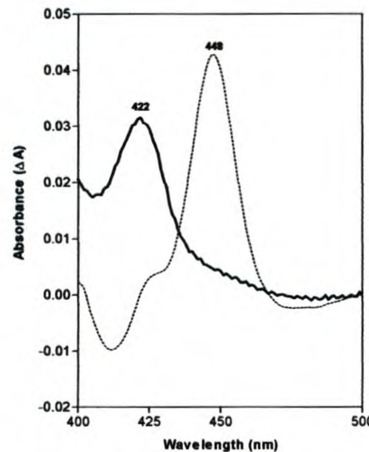

Figure 3.7 Reduced CO-induced induced difference spectra showing the stability of detergent solubilised proteins from *P. pastoris*. (·······) spectra of freshly prepared soluble fraction as seen in **Figure 3.6** A and (——) spectra of the same fraction after 2 months stored at –20°C.

The detergent-solubilised P45017α was unstable as indicated by the maximum absorbancy shift from 448 nm to 422 nm in a reduced CO-induced difference spectrum. The detergent-solubilised human P45017α with β-ME produced a type I substrate induced difference spectrum in the presence of 10 μM progesterone (Figure 3.8A) as well as 10 μM pregnenolone (Figure 3.8 B). The spin state change from low spin to high spin is characteristic of human P45017α.

Stellenbosch University http://scholar.sun.ac.za

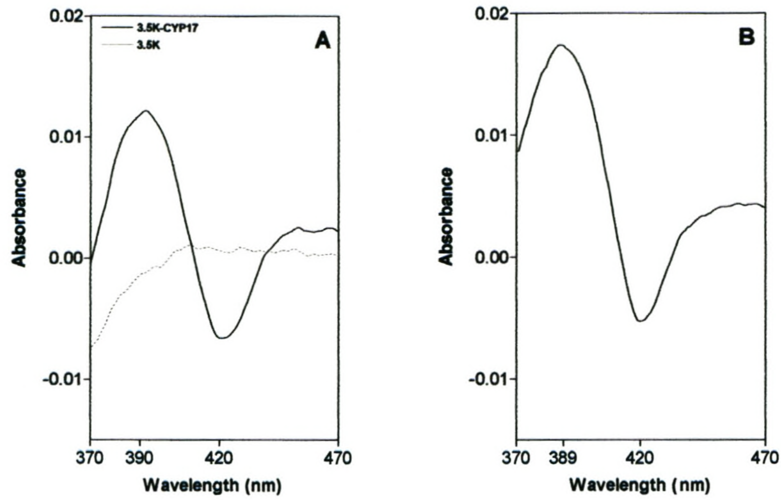

Figure 3.8 Substrate induced difference spectra of recombinant P45017α in a detergent-solubilised fraction (0.473 μM) in the presence of 10 μM progesterone (A) and 10 μM pregnenolone (B).

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 4 : THE DEVELOPMENT AND EVALUATION OF A PROTOCOL FOR THE PURIFICATION OF P45017α FROM MICROSOMES

4.1 Materials and Methods

4.1.1 Purification of Emulgen 913 solubilised P45017α

4.1.1.1 Ni^{2+}-NTA column chromatography

Ni-CAM™ HC affinity resin (5 ml suspension) (Sigma, St. Louis, USA) was washed with MQ and pre-equilibrated with equilibration/wash buffer (50 mM sodium phosphate, pH 8.0, 0.3 M NaCl, 20% v/v glycerol, 0.2% v/v Emulgen 913, 10 mM imidazole). Solubilised proteins (80 mg), isolated as described in section 3.1.2, was added at a concentration of 8.82 mg/ml to the equilibrated resin and gently agitated on a shaker (STR9, Stuart Scientific) for 19 hours at 4°C. A 5 × 0.9 cm perspex column was packed with the overnight incubated lysate-Ni^{2+}-NTA resin mixture resulting in a 2 ml bed volume (BV). During the packing of the column the flow-through was collected. The column was washed with wash buffer at a flow rate of 0.2 ml/min and the total volume was 20 ml. The histidine tagged P45017α was eluted from the column with elution buffer (50 mM sodium phosphate, pH 8.0, 0.3 M NaCl, 20% v/v glycerol, 0.2% v/v Emulgen 913, 250 mM imidazole) at a flow rate of 0.2 ml/min. After eluting ten fractions (2 ml each), the column was cleaned and regenerated according to the manufacturer's instructions. The collected fractions were stored at –20°C until analysed by SDS-PAGE and Western blot analyses.

4.1.1.2 Size exclusion chromatography

Eluate (300 μl) from the Ni^{2+}-NTA column, containing the histidine tagged P45017α protein was loaded on the Sephadex G-75 column (20 ml bed volume, flow rate 0.18 ml/min), which had been pre-equilibrated with loading buffer (100 mM sodium

Stellenbosch University http://scholar.sun.ac.za

phosphate, pH 7.6, 10% v/v glycerol). The fractions having an absorbance at 280 nm were collected and analysed by SDS-PAGE and Western blot analyses.

4.1.2 Purification of Emulgen 913 solubilised P45017α and direct column loading

Solubilised microsomal proteins (45 ml, 2.38 mg/ml), isolated as described in section 3.1.2, were loaded onto a Ni-CAM™ HC affinity resin (Sigma, St. Louis, USA) column (2 ml BV), which had been pre-equilibrated with equilibration buffer, at a flow rate of 0.2 ml/min. The column was washed with wash buffer (20 ml) containing imidazole (20 mM) and with wash buffer (25 ml) containing imidazole (50 mM) at a flow rate of 0.5 ml/min. Histidine tagged P45017α was eluted with elution buffer (12.5 ml) at a flow rate of 0.2 ml/min. The fractions with an absorbtion larger than 0.2 at 412 nm were pooled. An aliquot of the pooled fractions was analysed by SDS-PAGE while the rest was concentrated with an Amicon Ultracell-4 (Millipore). These preparations were also analysed with SDS-PAGE and Western blot analyses.

4.1.3 Purification of Emulgen 913/β-ME solubilised P45017α

Ni-CAM™ HC affinity resin (4 ml suspension) was washed with MQ and equilibrated with equilibration buffer containing Emulgen 913 (1% v/v). Solubilised microsomal proteins (90.8 mg), isolated as described in section 3.1.2.2, was added at a concentration of 4.54 mg/ml to the equilibrated resin and gently agitated on a shaker for 19 hours at 4°C. A 5 × 0.9 cm perspex column was packed with the overnight incubated lysate-Ni^{2+}-NTA resin mixture resulting in a 3 ml BV. During the column packing, the flow-through (37.5 ml) was collected. The column was washed with wash buffer (11 ml) containing Emulgen 913 (1% v/v) and imidazole (15 mM) and wash buffer (24 ml) containing Emulgen 913 (1% v/v) and imidazole (20 mM) at a flow rate of 0.15 ml/min. The histidine tagged P45017α was eluted from the column with elution buffer (4 ml) containing Emulgen 913 (1% v/v) at a flow rate of 0.15 ml/min. The column was cleaned and regenerated according to the manufacturer's instructions. The collected

Stellenbosch University http://scholar.sun.ac.za

fractions were pooled and stored at –20°C until analysed by SDS-PAGE and Western blot analyses.

4.1.4 Adsorption chromatography

Before removing the detergent from the Ni^{2+}-NTA purified P45017α fraction, a trial batch experiment was conducted to show that the hydrophobic polymeric Amberlite XAD-1600 (Rohm & Haas) resin was able to adsorb Emulgen 913.

An Emulgen 913 (1% v/v) solution was prepared and a spectrum was recorded between 500 and 200 nm with a Cary spectrophotometer (USA). The detergent solution (4 ml) was added to Amberlite XAD-1600 (Rohm & Haas) adsorption resin (4 ml BV) in a 10 ml Falcon tube. As a control deionised water (4 ml) was added to a second Falcon tube containing the same amount of resin. Both tubes were gently agitated on a shaker for 20 min at 4°C. The resins were left to settle and the top aqueous layers were analysed spectrophotometrically.

4.1.5 Two-step Purification of Emulgen 913/β-ME solubilised P45017α

Two methods for the preparation of pure human P45017α were evaluated. In the first method the detergent was first removed by adsorption chromatography followed by Ni^{2+}-NTA chromatography while in the second method the Ni^{2+}-NTA chromatography was carried out before detergent removal.

4.1.5.1 **Detergent removal with Amberlite XAD resin followed by Ni^{2+}-NTA column chromatography**

Two chromatography columns were packed. One, 12 × 0.9 cm, with Amberlite XAD-4 and one, 14 × 0.9 cm, with and XAD-1600. Both columns received identical treatment. After equilibrating the columns with buffer C (50 mM sodium phosphate, pH 7.4, 20% v/v glycerol and 20 mM NaCl) each was loaded with solubilised microsomal proteins (45 mg, 4.54 mg/ml, isolated as described in section 3.1.2.2). To ensure maximum adsorption of the detergent a flow rate of 0.15 ml/min was maintained. The fractions that had an absorbance greater than 0.2 at 412 nm were pooled. The XAD-4 column yielded 10 ml and the XAD-1600 column yielded 15 ml detergent free protein preparations.

Stellenbosch University http://scholar.sun.ac.za

Each Emulgen 913 free protein fraction was immediately added to a Ni-CAM™ HC affinity resin, previously equilibrated with equilibration buffer containing no Emulgen 913, and gently agitated on a shaker for 19 hours at 4°C. Two 5 × 0.9 cm perspex column were packed with the overnight incubated lysate-Ni^{2+}-NTA resin mixtures resulting in two 1 ml BV. During the packing the flow-throughs of both columns were collected (7 ml each). The columns were subsequently washed with wash buffer containing (15 mM imidazole), without Emulgen 913, at a flow rate of 0.1 ml/min and the total volume was 10 ml for each column. The histidine tagged P45017α was eluted from the columns with elution buffer without Emulgen 913 at a flow rate of 0.1 ml/min. After eluting five 2 ml fractions (10 ml total), the column was cleaned and regenerated according to the manufacturer's instructions. The collected fractions were concentrated 10 times with an Amicon Ultrafee-4 (Millipore) to a final protein concentration of 1.37 mg/ml and stored at –20°C until analysed by SDS-PAGE and Western blot analyses.

4.1.5.2 Ni^{2+}-NTA column chromatography followed by detergent removal with Amberlite XAD resin

Ni-CAM™ HC affinity resin slurry (Sigma, St. Louis, USA) (5 ml) was washed with MQ and equilibrated with equilibration buffer containing Emulgen 913 (0.1% v/v). Solubilised microsomal protein (136 mg), isolated as described in section 3.1.2.2, was added at a concentration of 4.33 mg/ml to the equilibrated resin and gently agitated on a shaker for 12 hours at 4°C. The lysate-Ni-NTA mixture was packed into a 5 × 0.9 cm perspex column (1.9 ml BV) and the flow-through was collected (50 ml). The column was washed with wash buffer (13.5 ml) containing Emulgen 913 (0.1% v/v) and imidazole (15 mM) at a flow rate of 0.15 ml/min. The histidine tagged P45017α was eluted from the column with elution buffer containing Emulgen 913 (0.1% v/v) and imidazole (100 mM) and elution buffer containing Emulgen 913 (0.1% v/v) and imidazole (250 mM) at a flow rate of 0.15 ml/min. The column was cleaned and regenerated according to the manufacturer's instructions.

The Ni^{2+}-NTA purified P45017α was added to equilibrated Amberlite XAD-1600 (5 ml BV) and gently agitated on a shaker for 20 min at 4°C. The Emulgen 913 was adsorbed

Stellenbosch University http://scholar.sun.ac.za

and the human P45017α was eluted in buffer C (20 ml). The eluted sample was concentrated 26 times in an Amicon Ultracell-4 and analysed by SDS-PAGE and Western blot analyses.

4.2 Results

4.2.1 Purification of P45017α with Emulgen 913 and incubation overnight

The first attempt to purify histidine-tagged P45017α from a detergent-solubilised yeast fraction (see Materials and Methods 4.1.1) with a Ni^{2+}-NTA column yielded a fairly pure P45017α protein at around 54 kDa (Figure 4.1, lanes 5-7). The majority of P45017α was in fraction 2, however two dominant proteins (81 and 42 kDa) co-eluted.

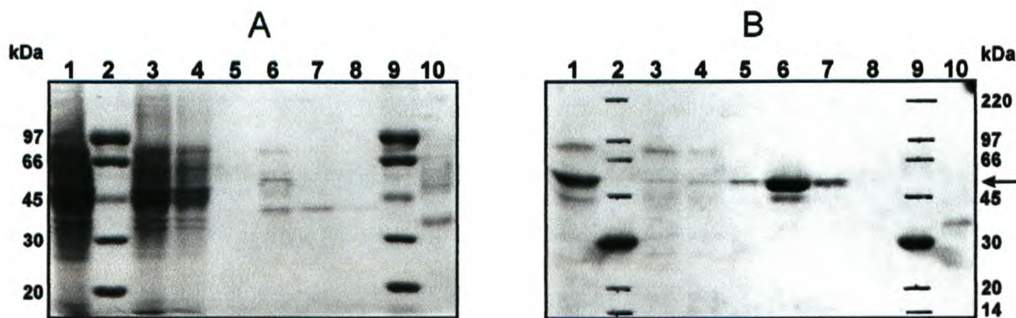

Figure 4.1 SDS-PAGE (A) and Western blot (B) analyses of Ni^{2+}-NTA purified proteins from the detergent-solubilised fraction. Lane 1: crude cell lysate (20 µg) loaded on column. **Lanes 2 and 9:** Low-molecular weight Marker. **Lane 3:** First fraction of flow through while packing the column. **Lane 4:** Last fraction of flow through while packing the column. **Lanes 5-9:** Fractions of eluting Ni-purified **P45017α. Lane 10:** Positive control containing a histidine-tagged ureate oxidase (10 µg). P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

An attempt to separate P45017α from these proteins by gel size exclusion chromatography with Sephadex G-75 was unsuccessful (Figure 4.2 and Figure 4.3 lane 4).

Stellenbosch University http://scholar.sun.ac.za

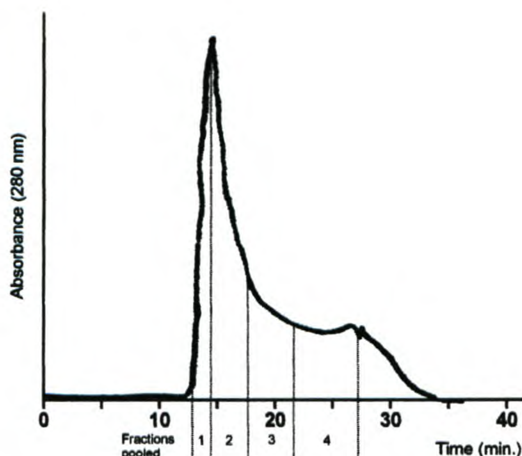

Figure 4.2 Chromatogram of the purification of P45017α with size exclusion chromatography. The eluting proteins detected at 280 nm were separated and pooled into four fractions.

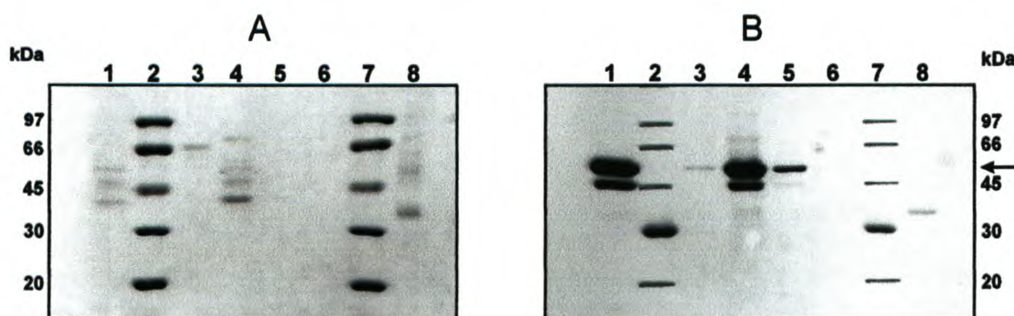

Figure 4.3 SDS-PAGE (A) and Western blot (B) analyses of gel filtration separated proteins. Lane 1: Ni^{2+}-NTA purified P45017α loaded on the G75 column. **Lanes 2 and 7:** Low-molecular weight Marker. **Lanes 3-6:** The 4 pooled fractions of the eluted proteins. **Lane 8:** Positive control containing a histidine-tagged ureate oxidase (10 μg). P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

To optimise the Ni^{2+}-NTA purification P45017α a direct purification procedure was carried out.

4.2.2 Purification of P45017α with Emulgen 913 and direct column loading

Direct loading of detergent-solubilised proteins, increased NaCl concentration (from 0.3 to 0.5 M) and step wise washing steps with increased imidazole concentration (20 to 50 mM imidazole) in the wash buffer (see more detail in Materials and Methods 4.1.2), did prevent non specific binding of contaminating proteins, but most of P45017α was eluted (Figure 4.4, lanes 3 and 4). The elution of P45017α at an imidazole concentration of 50

Stellenbosch University http://scholar.sun.ac.za

mM, indicates that the binding of P45017α to the resin is not high and the column must not be washed with imidazole concentrations higher than 15 mM. It was necessary to incubate the lysate with the resin overnight.

The concentrated fractions from the washing and elution steps show a prominent protein band at a molecular mass corresponding to 40 kDa in the SDS-PAGE shown in Figure 4.4C, lanes 3-8. This protein was not detected with the HisProbe in the Western blot analysis (Figure 4.4D, lanes 3-8) and was therefore not a histidine tagged or histidine rich protein. This unidentified protein can remain bound to the resin at very high imidazole concentrations (up to 250 mM). Another contaminating protein (45 kDa) co-eluted with the P45017α protein was also detected in the Western blot analysis (Figure 4.4D, lanes 3-5). This could be either a partially proteolytic degraded P45017α, a histidine rich protein or a metalloenzyme.

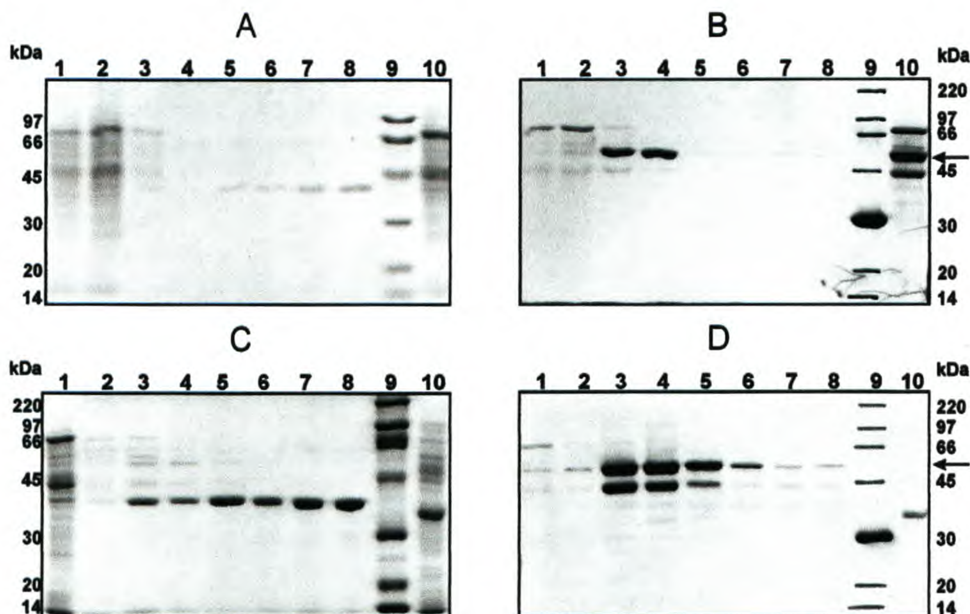

Figure 4.4 SDS-PAGE (A and C) and Western blot (B and D) analyses of Ni^{2+}-NTA purified proteins from the detergent-solubilised fraction. The protein fractions seen in SDS-PAGE A were concentrated with Amicon and run again on SDS-PAGE C. A and B: Lanes 1and 2: flow-through while packing the column. **Lane 3:** wash I step. **Lane 4:** wash II step. **Lanes 5-8:** elution of P45017α. **Lane 9:** Low-molecular weight Marker. **Lane 10:** crude soluble extract incubated with resin. **C and D: Lanes 1 and 2:** flow-through while packing the column. **Lane 3:** wash I step. **Lane 4:** wash II step. **Lanes 5-8:** elution of P45017α. **Lane 9:** Low-molecular weight Marker. **Lane 10:** Positive control containing a histidine-tagged ureate oxidase (10 μg). P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

Stellenbosch University http://scholar.sun.ac.za

It appeared that the 40 and 45 kDa proteins were always co-eluting with P45017α, hence a disulphide bond between the expressed P45017α and the contaminating proteins might have formed. The complex will be purified during the Ni^{2+}-NTA step and as soon as the samples were denatured by (β-mercaptoethanol) in the SDS loading buffer, the disulphide bonds were broken, thus two protein bands and P45017α appeared on the SDS-PAGE (Figure 4.4C, lanes 3-5).

The fraction shown in Figure 4.5 lane 3, was analysed by electro spray mass spectrometry (ESMS) to determine the molecular masses of the size proteins present in the preparation. ESMS results (Figure 4.5) revealed only a characteristic ion pattern of a polymer or detergent. This chromatogram is of Emulgen 913 in the sample which was not successfully removed during the desalting and concentration steps. Emulgen 913 has a very low CMC (critical miscellar concentration) therefore it is neither possible to separate nor dialyse the Emulgen 913 away from the protein sample.

4.2.3 Purification of Emulgen 913/β-ME solubilised P45017α

Two additional modifications to the purification protocol were made (see Materials and Methods 4.1.3) Firstly, β-ME was added to the solubilised cell lysate to prevent any disulfide bonds forming between P45017α and the contaminating proteins. A second purification step, to remove the Emulgen 913 from the P45017α sample, was also implemented. Hydroxyapatite was used successfully to remove the detergent (18, 7), however Amberlite XAD-1600 (Rohm & Haas) has not been shown previously to remove detergent from protein containing solution.

Stellenbosch University http://scholar.sun.ac.za

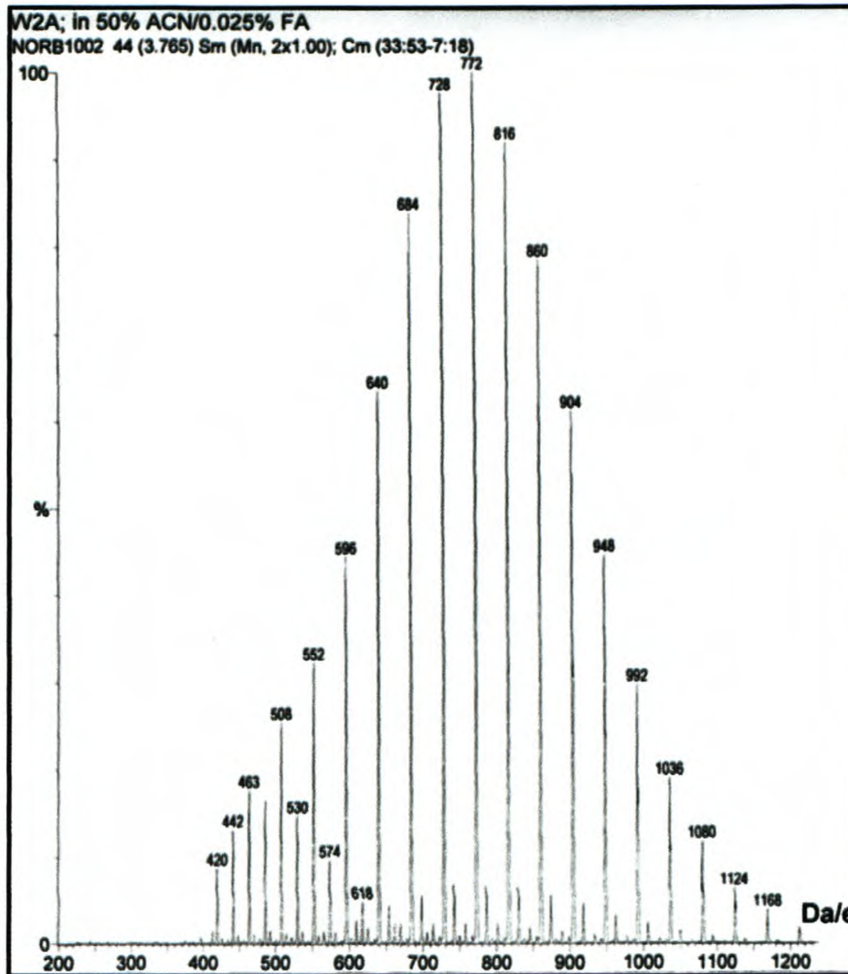

Figure 4.5 ESMS analyses of proteins in the wash buffer Figure 4.4 lane 3. Characteristic ion pattern of Emulgen 913.

The detergent solubilised lysate from *P. pastoris,* incubated overnight with the Ni^{2+}-NTA resin, was packed into a column. Washing was done in the presence of detergent by applying a step-wise gradient of imidazole and elution was achieved at high imidazole concentration. The elution of different proteins was monitored at 412 nm and recorded as seen in Figure 4.6. The chromatogram shows 2 distinct peaks, one during the flow-through step and one during the elution step. SDS-PAGE and Western blot analyses of these peaks show that the contaminating proteins (peak 1), which previously co-eluted with P45017α, were washed away Figure 4.7, lane 3, and a pure P45017α protein (peak 2) was eluted in Figure 4.7, lane 6.

Stellenbosch University http://scholar.sun.ac.za

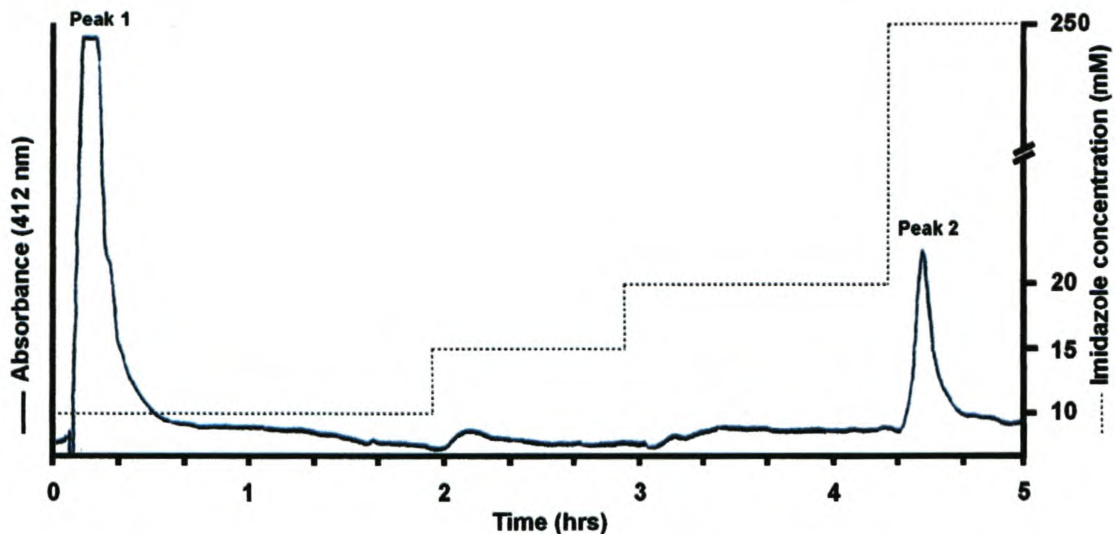

Figure 4.6 Elution profile of human P45017α purified on a Ni^{2+}-NTA column. Absorption at 412 nm (———); imidazole concentration (········).

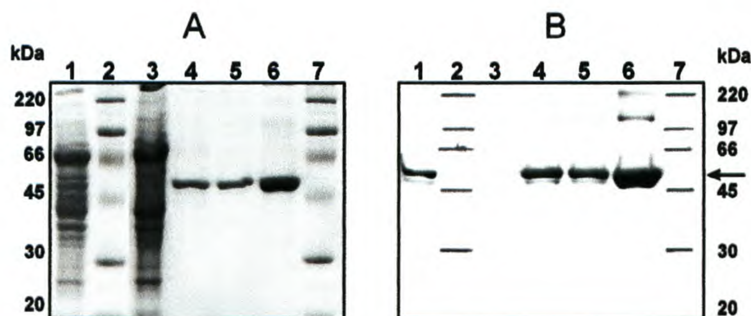

Figure 4.7 SDS-PAGE (A) and Western blot (B) analyses of proteins from the IMAC purification step: Lane 1: microsomal crude cell lysate (23 μg) loaded on Ni^{2+} column. **Lanes 2 and 7:** High-molecular weight Marker. **Lane 3:** flow-through (76 μg) during packing the column. **Lane 4:** wash I step (0.4 μg) (15 mM imidazole). **Lane 5:** wash II step (0.32 μg) (20 mM imidazole). **Lane 6:** elution (1.38 μg) of P45017α (250 mM imidazole). P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

Results further show that an electrophoretic more homogeneous P45017α fraction eluted in the washing steps (Figure 4.7 lanes 4 and 5). From these results it followed that the optimum elution for P45017α would probably be between 15-100 mM imidazole. Owaki A. *et al.* have shown that, during the purification with Ni^{2+}-NTA of the guinea pig P45017α, the protein was also eluted between 0-40 mM of imidazole (20). Therefore it can be concluded that the histidine tail incorporated into P450 enzymes has a relative weak binding affinity for the Ni^{2+}-NTA.

Stellenbosch University http://scholar.sun.ac.za

4.2.4 Adsorption chromatography

Purifying protein samples from Emulgen 913 containing solutions with XAD-4 and XAD-2 have previously been demonstrated (8). We carried out an experiment with XAD-1600 determine the efficiency of Emulgen 913 removal. Emulgen 913 has an aromatic ring structure which absorbs and therefore allows monitoring at 280 nm. Results show that Emulgen 913 was adsorbed by Amberlite XAD-1600, as seen in Figure 4.8.

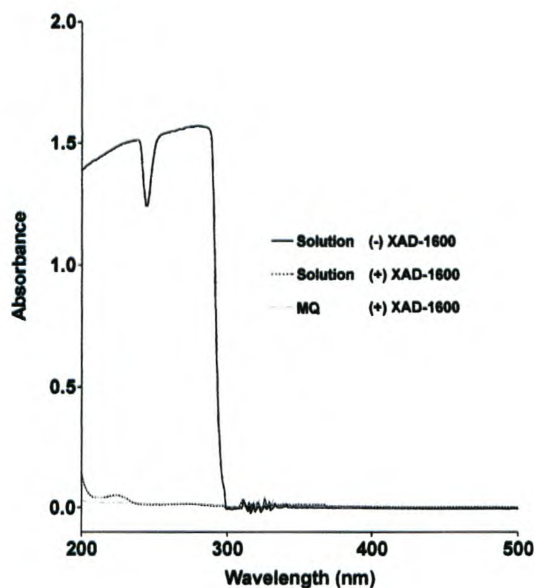

Figure 4.8 Wavelength scan of the adsorption of Emulgen 913 (1% v/v) solution by Amberlite XAD-1600.

4.2.5 Two-step Purification of Emulgen 913/β-ME solubilised P45017α

Two experiments (see Materials and Methods 4.1.5) were conducted to see which purification sequence will produce a pure P45017α. Either removing the detergent via adsorption chromatography first followed by Ni^{2+}-NTA chromatography or vice versa.

4.2.5.1 Detergent removal followed by Ni^{2+}-NTA chromatography

Two different adsorption resins (Amberlite XAD-4 and XAD-1600) were used to compare the efficiency of removing the Emulgen 913 from the crude solubilised microsomal lysate. Results show that P45017α was separated from Emulgen 913,

Stellenbosch University http://scholar.sun.ac.za

because the eluted protein fraction contained P45017α (Figure 4.9, lanes 3 and 4) but lacked the typical Emulgen 913 spectrum, indicating detergent removal by the resins. The XAD-1600 resin adsorbed proteins up to 40 kDa compared to XAD-4, which only adsorbed up to 20 kDa as seen in Figure 4.9, lanes 3 and 4. This was expected since the former has a bigger pore size. Although both type of resins offer the same high surface area and controlled pore size, XAD-1600 is a better resin to use due to the closely controlled particle size which results in better separation.

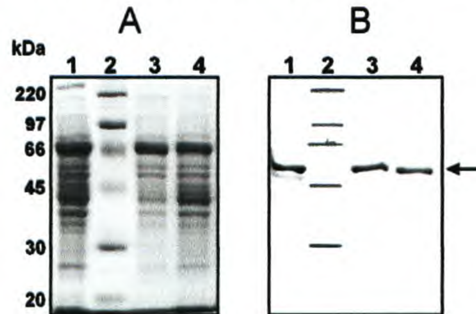

Figure 4.9 SDS-PAGE (A) and Western blot (B) analyses of purified P45017α with Adsorption resins. Lane 1: microsomal crude cell lysate loaded on XAD columns. **Lane 2:** Low molecular weight Marker. **Lane 3:** XAD-1600 purified proteins. **Lane 4:** XAD-4 purified proteins. P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

The detergent-free P45017α was subsequently purified with Ni^{2+}-NTA and results show that the P45017α could not be purified to electrophoretic homogeneity (see Figure 4.10, lanes 5 and 7). The co-elution of a 40 kDa protein was again present as was seen in Figure 4.10, lanes 3-8.

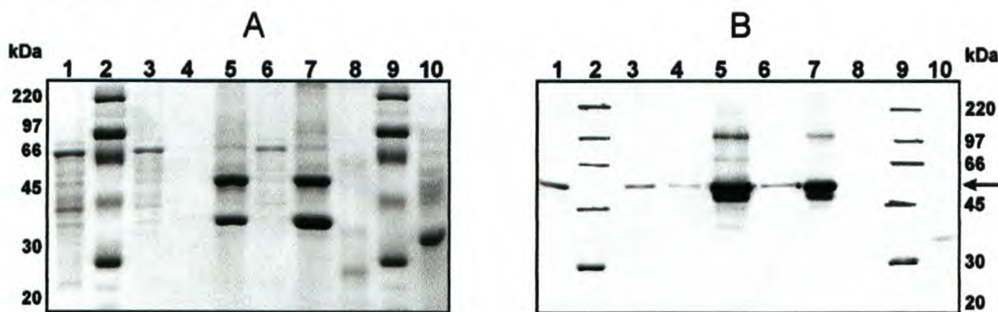

Figure 4.10 SDS-PAGE (A) and Western blot (B) analyses of the 2-step purification of P45017α: Lane 1: microsomal crude cell lysate loaded on XAD columns. **Lanes 2 and 9:** Low molecular weight Marker. **Lanes 3-5:** Proteins that were first purified by XAD-4 and subsequently loaded onto Ni^{2+}-NTA column, lanes 3-5 is the flow-through, wash and elution step from the Ni^{2+}-NTA column, respectively. **Lanes 6-8:** Proteins that were first purified by XAD-1600 and subsequently loaded onto Ni^{2+}-NTA column, lanes 6-8 is the flow-through, wash and elution steps from the Ni^{2+}-NTA column, respectively. **Lane 10:** Positive control containing a histidine-tagged ureate oxidase (10 μg). P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

Stellenbosch University http://scholar.sun.ac.za

4.2.5.2 Ni^{2+}-NTA chromatography followed by detergent removal

The detergent solubilised lysate from *P. pastoris,* incubated overnight with the Ni^{2+}-NTA resin, was packed into a column. Washing was carried out in the presence of detergent by applying a step-wise gradient of imidazole and elution was achieved at high imidazole concentration (see Material and Methods 4.1.5.2). The elution of different proteins was monitored at 412 nm and the recorded chromatogram is shown in Figure 4.11. Two distinct peaks were detected. Peak 1 indicated elution of material during the flow-through step and Peak 2 indicated elution of material during elution. SDS-PAGE and Western blot analyses of these peaks show that the contaminating proteins (peak 1), which previously co-eluted with P450$17\alpha$ (Figure 4.12, lane 5), were effectively separated from P450$17\alpha$ which eluted with 100 mM imidazole buffer (Figure 4.12, lane 7) There was no detectable of P450$17\alpha$ in the flow-through fraction of the Ni^{2+}-NTA column (Figure 4.12, lane 5), which indicated complete binding of P450$17\alpha$ without any overloading of the resin. The optimal concentration of imidazole in the elution buffer was 100 mM which was empirically determined from previous purification experiments. No detectable P450$17\alpha$ was present in the 250 mM imidazole-eluting fraction (Figure 4.12, lane 8), indicating that 100 mM imizazole was sufficient to remove all the P450$17\alpha$ from the column. The XAD-1600 adsorption resin produced a detergent free P450$17\alpha$ which migrated as a single band on SDS-PAGE with an estimated molecular mass of 58000 Da (Figure 4.12, lane 10). ES-MS analyses showed that the detergent was removed successfully by XAD-1600 resin but the protein concentration was too low for a protein mass spectrum and the volume obtained in this preparation was too small to record a reduced CO-induced P450 difference spectrum (results not shown).

Stellenbosch University http://scholar.sun.ac.za

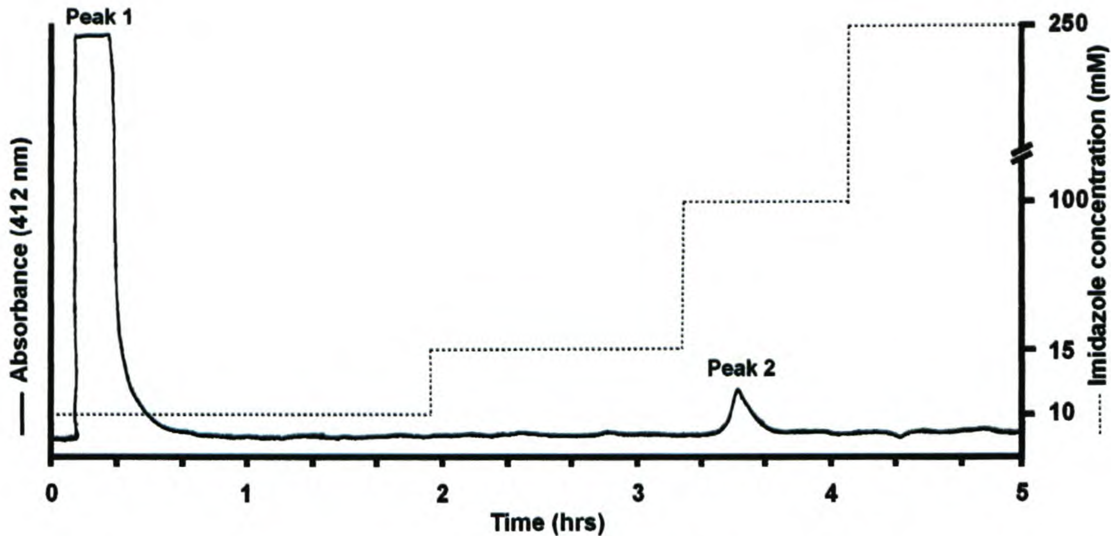

Figure 4.11 Elution profile of human P45017α purified on a Ni^{2+}-NTA column. Absorption at 412 nm (———); imidazole concentration (·········).

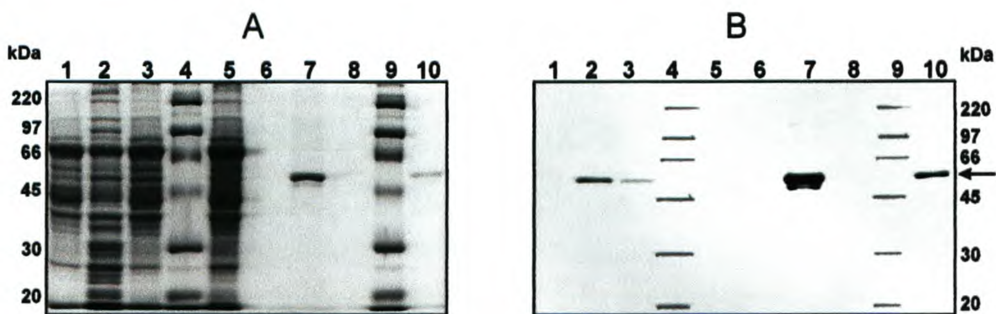

Figure 4.12 SDS-PAGE (A) and Western blot (B) analyses of the 2-step purification of P45017α: Lane 1. detergent-solubilised proteins from GS115/3.5K cells. **Lanes 2 and 3:** insoluble microsomal membranes and detergent-solubilised proteins from GS115/3.5K-CYP17 cells, respectively. The latter was loaded on the Ni^{2+}-NTA column. **Lanes 4 and 9:** Low molecular weight Marker. **Lane 5:** The flow-through and wash I step from the Ni^{2+}-NTA column. **Lane 6:** wash II step. **Lane 7:** elution of P45017α with imidazole (100 mM). **Lane 8:**elution of proteins with imidazole (250 mM). **Lane 10:** Purified P45017α obtained by Ni^{2+}-NTA chromatography followed by detergent removal with XAD-1600 resin. P45017α (58 kDa) is indicated by the arrow. Detection was carried out using an INDIA HisProbe-HRP at 1:5000.

A cytochrome P450 CO-induced difference spectrum Figure 4.13 with an absorption maximum of 422 nm was recorded from the Ni^{2+}-NTA purified P45017α Figure 4.12, lane 7, which had been kept for 2 months at –80°C. Results show that the protein had denatured to P420 nm, indicating an expected denaturation since the sample was dissolved in the elution buffer in the presence of 250 mM imidazole. The imidazole acts as the fifth ligand to the active site of P45017α forming a charge transfer complex.

Stellenbosch University http://scholar.sun.ac.za

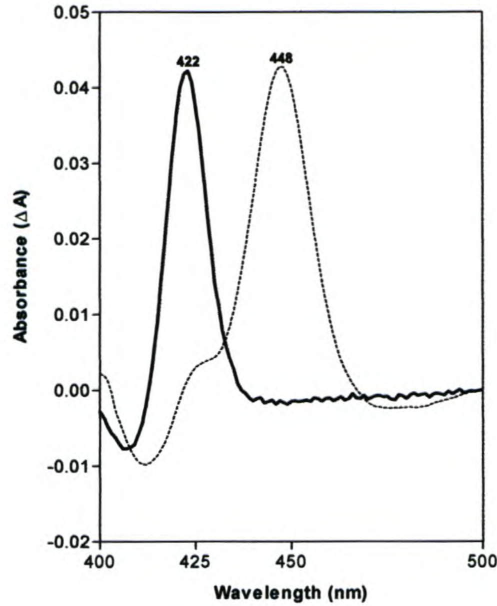

Figure 4.13 Reduced CO-induced induced difference spectra of Ni^{2+}-NTA purified P450$17\alpha$; (·······), Detergent-solubilised protein fraction from GS115/pPIC3.5K-CYP17 cells loaded onto the Ni^{2+}-NTA column; (——) Ni^{2+}-NTA purified P450$17\alpha$. The latter sample suspended in (50 mM sodium phosphate, pH 8.0, 0.3 M NaCl, 20% v/v glycerol, 0.1% v/v Emulgen 913, 250 mM imidazole) was diluted 1:1 in 100 mM sodium phosphate buffer, pH 7.0 which was previously saturated with carbon monoxide.

The isolation and purification of P450$17\alpha$ expressed by *P. pastoris* is summarised in Table 4.1.

Table 4.1 Purification of P450$17\alpha$ from a one liter *P. pastoris* cell culture.

Purification steps	Protein conc. (mg/ml)	Total Volume (ml)	Total protein (mg)	P450 conc. (μM or nmol/ml)[a]	Total P450 (nmol/L culture)	Specific content. (nmol/mg protein	Yield (%)
Microsomes	6.70	5	33.5	0.703	3.5	0.104	
Clear solubilised lysate[b]	4.54	50	227	0.946	47.3	0.208	100
Ni^{2+}-NTA eluate	0.52	30	15.6	0.928	27.8	1.780	58.7
XAD-1600 eluate	nd[c]			nd			

[a] The total amount of P450 was determined by the reduced CO-induced difference spectra (3).

[b] The detergent solubilised cell lysate from *P. pastoris* was used for the purification.

[c] Not determined

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 5 : DISCUSSION

The awareness of the importance of steroidogenic human cytochrome P450 enzymes has increased significantly in recent years. The cDNA sequences for more than thirty different CYP17's have been deposited in databases (e.g. GenBank)[1]. In this context human cytochrome P45017α can be singled out as one of the steroidogenic cytochromes P450 which attracted the attention of many researchers for a number of reasons including its role in human disease and its unique dual hydroxylase and lyase activities. The enzyme is also currently being examined for its ability to catalyse the formation of 16-ene steroids, via 16-ene-synthase activity (9) and has also lately been indicated as an important factor involved in polycystic ovarian syndrome (PCOS) (10, 11) and pheromone biosynthesis. The interest shown in this microsomal P450 enzyme in these current studies requires the demand for sufficient quantities of pure human P45017α protein. The availability of reasonable amounts of human P45017α will facilitate studies of this fascinating steroidogenic hemoprotein and will go a long way in facilitating our search for the functional characteristics of this enzyme. In addition, sufficient amounts of a pure P45017α can also be used to raise antibodies against P45017α, which are currently not readily available. Moral issues, aggregation in the absence of detergents, relative low concentrations and instability have severely complicated or hamperd efforts to isolate and purify P45017α from human tissue.

Heterologous expression systems have been used to produce relatively high concentrations of cytochrome P45017α proteins from various vertebrates, such as bovine (12), rat (13), pig (14), hamster (15), trout (16) and shark (17), but the human P45017α enzyme has only been overexpressed for purification and characterisation in *E. coli* (18). However, modifications of the amino terminal sequence of human P45017α, precisely described for bovine P45017α by Barnes *et al.* (12) resulted in high-level expression in *E. coli*. This modification also increased expression levels of the shark P45017α (19, 20). For expression in prokaryotes alterations within seven codons of the cDNA, as described by Barnes (21), created an mRNA that is efficiently translated and minimises secondary structure formation (22).

---

[1] Prof R W. Estabrook, personal communication.

Stellenbosch University http://scholar.sun.ac.za

During this investigation the human cytochrome P45017α was successfully cloned and overexpressed P45017α in the methylotrophic yeast, *P. pastoris* (23, 24), without modifying the 5' terminal of CYP17 cDNA. *P. pastoris* has become a very attractive system for heterologous expression of heme-containing proteins (25, 26, 27, 28) and three P450-specific enzymes from spiny dogfish shark(29), the spiny lobster (30) and Cassava plant (31) (CYP17, CYP2L1 and CYP71D1, respectively) have functionally been expressed in this organism which suggests that *P. pastoris* has the potential capability of producing active heme-containing proteins. In our hands the expression level of human P45017α in the *P. pastoris* system was 50 nmol of P45017α/L of culture, as measured by P45017α recovered in the solubilised lysate. This figure was slightly lower than results obtained for bovine P45017α previously expressed in baculovirus (32). The total cytochrome P450 content was 300 pmol/mg microsomal protein, which is significantly higher when compared to the *P. pastoris* expressed shark P45017α (15-30 pmol/mg microsomal protein), lobster CYP2L1 (20 pmol/mg microsomal protein) and plant CYP79D1 (125 pmol/mg microsomal protein). The low expression of shark P45017α in *P. pastoris* prevented Trant to determine the cytochrome P450 content of whole cells, lysed cells or microsomal fraction from *P. pastoris* (29). The high expression levels of human P45017α may either be due to the optimised context of the translational initiation consensus sequence or of a gene dosage effect. Multi-copy integration of the CYP17 into the yeast genome may have occurred, since the clones have been selected on 0.5 mg/ml G418 antibiotics. The copy number was, however, not determined by Southern blot analysis.

Growth of GS115 (Mut^+) cell cultures, during growth and induction phases, was normal and after 72 hours the cells reached cell densities with an OD_{600} = 12. The KM71 (Mut^s) cell cultures ceased growing abruptly after 48 h of induction. The two most likely explanations for this phenomenon could be: (i) there could have been a problem with the media and the cultures nutrients were limited; (ii) the cultures could have been contaminated at the 48 hours time point, because at that time the supplementation of methanol for induction was administered. The risks of contamination at these induction times was high, since the flasks had to be removed from the shaking incubators and the

Stellenbosch University http://scholar.sun.ac.za

cheese cloth removed to pipette the appropriate amount of methanol into the flasks. The control strain GS115/Albumin (Mut^{s}) also stopped growing after 48 hours, an indication that the medium was most likely contaminated since in the GS115 growth experiment their growth was normal. Therefore it was decided to continue with the GS115 strains in the subsequent expression and purification experiments.

P45017α was functionally expressed in *P. pastoris* as evidenced by the ability of recombinant yeast cells to convert progesterone to 17α-hydroxyprogesterone and 16α-hydroxyprogesterone. The conversion of progesterone in intact *P. pastoris* cells demonstrated that the endogenous yeasts reductase system was able to sustain electron transport to the human P45017α. At a total concentration of 10 μM, only 65% of the progesterone was converted within the first 6 min. The incomplete conversion of progesterone to its intermediates was probably caused either by the high initial concentration of progesterone, because activity studies in whole cells are usually executed with 1 μM of substrate (29, 32) or the cell density was too high (1×10^9 cells/ml) (33). The initial conversion rate was so high that the demand for reducing equivalents could not be supplied by the cells and were exhausted, leading to a decrease in conversion rate. The initial rate of the 17α-hydroxylation of progesterone (10 μM) in a 48 hr yeast culture expressing heterologous human P45017α, was 1.37 nmol/min/10^9 cells. This rate was 90 times faster than the shark P45017α expressed in *P. pastoris* (29). Another explanation for the rapid ceasure of the reaction after 5 min could be due to the limitation of oxygen required for the hydroxylation reaction. The assay was performed in small unbaffled Erlenmeyer flasks and shaken slowly, most probably resulting in an inefficient oxygen supply. The linear growth of the yeast cells, opposed to the logarithmic growth, substantiates the explanation of a lack of oxygen. This hypothesis is currently under investigation, however, to ensure a constant supply of oxygen during the assay is technically difficult and remains to be elucidated.

The solubilisation of P45017α from yeast microsomes with different detergents (Emulgen 913, Triton X-100), with or without sonication, was not very successful. Sonication was probably creating strong miscellar superstructures, which were holding P45017α so tightly bound that the detergents were not able to solubilise the protein from

Stellenbosch University http://scholar.sun.ac.za

these miscells. We were able to solubilise the majority of P45017α from the membrane fraction with drop-wise addition of a Emulgen 913 (1% v/v) to stirring cell lysate for 4 hours. The recombinant P45017α in the solubilised lysate and insoluble microsomal fraction produces a reduced cytochrome P450 CO-induced difference spectrum and type I substrate-induced difference spectra in the presence of progesterone and pregnenolone, indicating that the *P. pastoris* expression system produced a correctly folded hemeprotein. The spectrally detectable P45017α showed a peak at 448 nm and no peak was detected for the control strain (empty vector). This confirmed that, although the yeast has its own cytochrome P450 and OR responsible for ergosterol biosynthesis, the endogenous concentrations of these enzymes were low and had no effect on the expression or detection of human P45017α (34, 35).

The reduced CO-induced different spectra of the solubilised cellular protein fraction, isolated with two different buffers (buffer A and buffer BB), showed that in the presence of β-ME (10 mM), one dominant 448 nm peak with a shoulder at 427 nm occurred, compared to two peaks at 448 and 427 nm in the absence of β-ME. The 427 nm peak originates from mitochondrial cytochrome oxidase (absorption maximum around 430 nm) (36, 37) and is not native b_5 or cytochrome b1, since it cannot be reduced with NADPH (38). The presence of β-ME most probably reduced the mitochondrial cytochrome oxidase resulting in a low 427 nm peak.

The isolation of the membrane fraction of *P. pastoris* was omitted and only the solubilised cell lysate was used in the purification procedures. To purify recombinant human P45017α from *P. pastoris*, we have employed a quick 2-step purification procedure consisting of metal chelate affinity chromatography (Ni^{2+}-NTA) followed by adsorption chromatography. The decision was made to incorporate a six histidine tag at the carboxyl terminus, because previous reports have shown that the presence of the affinity tag at the carboxyl terminus did not significantly influence the activity of the purified cytochromes P450 (18, 39, 40)

An overnight incubation was required between the detergent-solubilised yeast fraction and the Ni^{2+}-NTA resin to ensure proper binding of the histidine-tagged P45017α to the

resin. With direct loading of the P45017α onto the Ni^{2+}-NTA column the protein eluted with the washing buffer. During Ni^{2+}-NTA purification it is crucial that the crude microsomal extract has β-mercaptoethanol present, since β-ME prevents disulphide bridges between cysteine residues of P45017α and other contaminant proteins (81 and 42 kDa), thereby preventing co-aggregation. The unwanted proteins can then be readily eluted in the wash step. This binding of native contaminants to recombinant expressed protein via disulfide bonds has been regularly observed in *E. coli* expression systems (41). Dithiotreitol (DTT) is normally added to all the buffers when working with P450 enzymes. DTT, however, reduces the nickel ions of the Ni^{2+}-NTA resin and therefore DTT was replaced with of β-ME (10 mM) in the solubilisation buffer.

Other contaminants interfering during IMAC purification are endogenous histidine-rich or metal-binding proteins from *P. pastoris* which interfere with the histidine tag of P45017α by co-purifying or outcompeting recombinant histidine-tagged P45017α when using a Ni^{2+}-NTA column. A 100 kDa protein co-eluted with P45017α in the elution buffer containing imidazole (250 mM) and a signal was observed in the Western blot analysis. However, when using 100 mM of imidazole, only P45017α (i.e. weaker binding for nickel) was eluted, but the bigger size protein was not (i.e. stronger affinity for nickel). Therefore the optimum elution range of P45017α was between 15-100 mM imidazole. Interestingly, research has shown that generally low imidazole concentrations were used for eluting histidine-tagged cytochromes P450 during IMAC purification. (42, 20). Schalk *et al.* found that his CYP73A1 with four histidine residues at carboxyl terminal eluted with 50 mM imidazole (39) while Kempf *et al.* used 80 mM imidazole to elute the CYP2D6 with a six histidine tag at the amino terminal (40).

Emulgen 913 in Ni^{2+}-NTA purified P45017α fraction was removed successfully with a polymeric adsorbant Amberlite XAD-1600. Although XAD-2 and XAD-4 resins have been shown to remove detergents from P450 enzyme fractions (43), XAD-2 has been discontinued by the Rohm and Haas Company. The newly developed XAD-1600 resin yielded a finer separation than XAD-4 by combining narrow particle distribution, good mesoporosity and high surface area.

Stellenbosch University http://scholar.sun.ac.za

The investigation of the sequence between the two purification steps, i. e. first using Ni-NTA chromatography and subsequently using the adsorption chromatography or vica versa, has been considered. Results show that removing Emulgen 913 and β-ME with adsorption chromatography, prior to Ni^{2+}-NTA purification, creates a nonreduced environment allowing for the aggregation of the 40 kDa protein with P45017α, which co-elutes in the Ni^{2+}-NTA elution step.

The CO induced difference spectrum of reduced electrophoretically pure P45017α protein, obtained from the Ni-NTA chromatography, showed an absorbance at 420 nm. To prevent this degradation of the P450 enzyme, the addition of a ligand, including an inhibitory ligand, to the binding site of the specific P450 protein is necessary as shown in literature (44, 45). Adding a ligand like progesterone to the buffers not only stabilises the protein, but it specifically protects the membrane protein from proteolytic digestion (46), and purification with little conversion to P420 was observed (47, 48). The CO different spectrum of the reduced purified bovine P45017α under detergent-solubilised conditions was not detectable in the absence of progesterone, indicating the extreme instability of the enzyme in the substrate free form (34). Kominami *et al.* has shown that the stability of purified bovine P45017α without progesterone can only be maintained in liposome membranes and similar stabilisations of membrane proteins in liposome membranes have been shown (49).

The *P. pastoris* expression system for steroidogenic cytochromes P450 opens possibilities for future projects. The conversion of pregnenolone in intact yeast cells and activity assays of progesterone and pregnenolone in microsomal fraction warrant further investigation. In addition to the presently developed expression systems studying mammalian cytochrome P450 (50, 51), this genetically modified Pichia expression system can assist to study the activities of expressed cytochrome P450 in the presence of inhibitors and reducing equivalents in living cells or in the microsomal milieu. To fully understand the cellular environment of the unexplored *P. pastoris*, the P450 content, endogenous b_5, OR and cytochrome oxidase should be investigated as all these accessory proteins might impact on the heterologously expressed steroidogenic cytochromes P450. The *P. pastoris* expression system for steroidogenic cytochromes P450 can also be scaled up from shake-flasks to fermentors to further increase the yield

Stellenbosch University http://scholar.sun.ac.za

of P450 proteins. Proper enzyme kinetic studies can be conducted in bioreactors as cells can be grown under well controlled conditions. Further increase in yield can be accomplished by producing multi-copy strains with the pPIC3.5K-CYP17 vector.

The production of polyclonal and monoclonal antibodies against the purified human P45017α obtained in this study is currently under investigation. These antibodies will be used in studies of the developmental and regulatory features of human P45017α. The production of therapeutic inhibitors linked to these antibodies will also be investigated. P45017α is a target in pharmacological interest for prostatic androgen dependent diseases, such as begnin hypertrophy and carcinoma.

In conclusion, this human P45017α construct, overexpressed in *P. pastoris*, yielded sufficient amounts of functional P45017α. The two simple purification steps yielded sufficient amount of pure human P45017α for further studies on the enzyme which would include enzymology, control of catalysis, crystallisation and antibody production.

REFERENCES

1. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2. Laemmli, U. K. (1970) *Nature* **227**, 680-685.
3. Omura, T. and Sato, R. (1964) *J. Biol. Chem.* **239**, 2370-2378.
4. Yun, C.-H., Miller, G. P. and Guengerich, F. P. (2000) *Biochemistry* **39**, 11319-11329.
5. Pan, Z., Camara, B., Gardner, H. W. and Backhaus, R. A.(1998) *J. Biol. Chem.* **273**, 18139-18145.
6. Anderson, M. D., Busk, P. K., Svendsen, I. and Møller, B. L. (2000) *J. Biol. Chem.* **275**, 1966-1975.
7. Guengerich, F. P., Hosea, N. A. and Martin, M. V. (1998) in *Cytochrome P450 Protocolls* (I. R. Phillips and E. A. Shephard, ed.) pp. 77-84, Vol. 107, Humana Press Inc., Totawa, NJ.
8. Guengerich, F. P. and Martin, M. V. (1998) in Cytochrome P450 Protocolls (I. R. Phillips and E. A. Shephard, ed.) pp. 35-54, Vol. 107, Humana Press Inc., Totawa, NJ.
9. Soucy, P., Lacoste, L. and Luu-The, V. (2003) *Eur. J. Biochem.* **270**, 1349-1355.
10. Qin, K. and Rosenfield, R. L. (1998) *Mol. Cell. Endocrinol.* **145**, 111-121.
11. Auchus, R. J., Geller, D. H., Lee, T. C. and Miller, W. L. (1998) *Trends Endocrin. Met.* **9**, 47-50.
12. Barnes, H. J., Arlotto, M. P. and Waterman, M. R. (1991*) Proc. Natl. Acad. Sci. USA* **88**, 5597-5601.
13. Fevold, H. R., Lorence, M. C., McCarthy, J. L., Trant, J. M., Kagimoto, M., Waterman, M. R. and Mason, J. I. (1989) *Mol. Endocrinol.* **3**, 968-975.
14. Nakajin, S., Shinoda, M., Haniu, M., Shively, J. E. and Hall, P. F. (1984) *J. Biol. Chem.* **259**, 3971-3976.
15. Cloutier, M., Fleury, A., Courtemanche, J., Ducharme, L., Mason, J. I. and LeHoux, J.-G. (1995) *Ann. N. Y. Acad. Sci.* **774**, 294-296.
16. Sakai, Y., Tanaka, M., Avdachi, S., Miller, W. L. and Nagahama, Y. (1992) *FEBS Lett.* **301**, 60-64.
17. Trant. J. M. (1995) *Gen. Comp. Endocr.* **102**, 173-182.
18. Imai, T., Globerman, H., Gertner, J., Kagawa, N. and Waterman M. R. (1993) *J. Biol. Chem.* **268**, 17317-17375.
19. Trant, J. M. (1996) *Gen. Comp. Endocr.* **102**, 173-182.
20. Owaki, A., Takamasa, A., Yamazaki, T and Kominami, S. (2002) *J. Steroid Biochem.* **81**, 255-262.

Stellenbosch University http://scholar.sun.ac.za

21. Barnes, H. J. (1996) *Method. Enzymol.* **272**, 3-14.

22. Gold, L. and Stormo, G. D., (1987) In: "*Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology" (F. C. Neidhardt and J. L. Ingraham, eds.) pp.1302-1307. American Society for Microbiology, Washington, DC.

23. Cregg, J. M., Vedcick, T. S. and Raschke W. C. (1993) *Biotechnology NY* **11**, 905-910.

24. Higgins, D. R. and Cregg, J. M. (1998) In: *Pichia protocols* (D. R. Higgins and J. M. Cregg, eds.) Humana Press, Inc., Totawa, NJ.

25. Gachhui, R., Presta, A., Bentley, D. F., Abu-Soud, H. M., McArthur, R., Brudvig, G., Ghosh, D. K. and Stuehr, D. J. (1996) *J. Biol. Chem.* **271**, 20594-20602.

26. Leber, A., Hemmens, B., Klösch, B., Goessler, W., Raber, G., Mayer, B. and Schmidt, K. (1999) *J. Biol. Chem.* **274**, 37658-37664.

27. Skipper, L., Campbell, W. H., Mertens, J. A. and Lowe, D. J. (2001) *J. Biol. Chem.* **276**, 26995-27002.

28. Gu, L., Lajoie, C. and Kelly, C. (2003) *Biotechnol. Progr.*19, 1403-1409.

29. Trant, J. M. (1996) *Arch. Biochem. Biophys.* **326**, 8-14.

30. Boyle, S. M., Popp, M. P., Smith, W. C., Greenberg, R. M. and James, M. O. (1998) *Mar. Environ. Res.* **46**, 25-28.

31. Anderson, M. D., Busk, P.K., Svendsen, I and Møller, B. L. (2000) *J. Biol. Chem.* **275**, 1966-1975.

32. Barnes, H.J., Jenkins, C.M. and Waterman, M.R. (1994) *Arch. Biochem. Biophys.* **315**, 489-494.

33. Sakaki, T., Shibata, M., Yabusaki, Y., Murakami, H. and Ohkawa, H. (1989) *DNA Cell Biol.* **8**, 409-418.

34. Yabusaki Y. (1998) in *Cytochrome P450 Protocolls* (I. R. Phillips and E. A. Shephard, ed.) pp. 195-202, Vol. 107, Humana Press Inc., Totawa, NJ.

35. van den Brink, H. J. M., van Gorcom, R. F. M., van den Hondel, C. A. M. J. J. and Punt, P. J. (1998) *Fungal Genet. Biol.* **23**, 1-17.

36. Käppeli, O., Sauer, M. and Fiechter, A. (1982) *Anal. Biochem.* **126**, 179-182.

37. Käppeli, O. (1986) *Microbiol. Rev.* **50**, 244-258.

38. Ishidate, K., Kawaguchi, K., Tagawa, K. and Hagihara, B. (1969) *J. Biochem.* **65**, 375-383

39. Schalk, M., Nedelkina, S., Schoch, G., Batard, Y. and Werck-Reichhart, D. (1999) *Biochemistry* **38**, 6093-6103.

40. Kempf, A. C., Zanger, U. M. and Meyer, U. A. (1995) *Arch. Biochem. Biophys.* **321**, 277-288.

41. Wulfing, C., Lombardero, J. and Pluckthun, A. (1994) *J. Biol. Chem.* **269**, 2895-2901.

42. Yanazaki, T., Ohno, T., Sakaki, T., Akiyoshi-Shibata, M., Yabusaki, Y., Imai, T. and Kominami, S. (1998) *Biochemistry* **37**, 2800-2806.

43. Bhamre, S., Anandatheerathavarada, H. K., Shankar, S. K., Boyd, M. R. and Ravindranath, V. (1993) *Arch. Biochem. Biophys.* **301**, 251-255.

44. Larson, J.R., Coon, M.J. and Porter, T.D. (1991) *J. Biol. Chem.* **266**, 7321-7324.

45. Sandhu, P., Guo, Z., Baba, T., Martin, M. V., Tukey, R. H. and Guengerich, F. P. (1994) *Arch. Biochem. Biophys.* **309**, 168-177.

46. Kühn-Velten, W. N. and Löhr, J. B. (1996) *FEBS Lett.* **388**, 21-25.

47. Kominami, S., Shinzawa, K. and Takemori, S. (1982) *Biochem. Biophys. Res. Commun.* **109**, 916-921.

48. Suhara, K., Fujimura, Y., Shiroo, M. and Katagiri, M. (1984) *J. Biol. Chem.* **259**, 8729-8736.

49. Kominami, S., Higuchi, A. and Takemori, S. (1988) *Biochim. Biophys. Acta* **937**, 177-183.

50. Auchus, R. J., Lee, T. C. and Miller, W. L. (1998) *J. Biol. Chem.* **273**, 3158-3165.

51. Ehmer, P. B., Bureick, M., Bernhardt, R., Müller, U. and Hartmann, R. W. (2002) *J. Steroid Biochem.* **81**, 173-179.

CONCLUSION

Insertion of a foreign gene into an expression vector does not guarantee a high level of expression of the foreign protein. Heterologous expression is a complex multi-step process and problems can arise at numerous stages, from transcription through to protein stability. Results dealing with failures and even low yields of heterologously expressed P450 enzymes are often not reported and has hindered biotechnologists to readily produce a sought after protein. Wrong options with regards to hosts, promoters and coding sequences are frequently chosen and novel ideas or strategies to improve expression levels are difficult to formulate. However, with this study, there is now a considerable accumulated experience on P450 gene expression, especially in yeasts. This overview can be informative and helpful in the identification of a particular problem during expression and may contribute to quicker solutions. We have investigated all the different expression systems that were used and focused on those systems which are the most suitable for the large-scale production of expressed cytochromes P450.

Not only was the main objective, that of producing a pure unmodified P450 enzyme achieved, but *P. pastoris* was shown to be an excellent host, as was demonstrated with the human P45017α expression. The success of overexpressing and producing unmodified human P450 enzymes in *P. pastoris* offers several potential applications in P450 research. This study was merely a stepping stone for future projects such as — development of activity assays in intact cells, large-scale P450 production with alternative reducing sources, multi-copy expressions and industrial fermentations.

Undoubtedly, *P. pastoris* will be the next generation of heterologous expression system used as a tool to elucidate and solve many debated issues regarding the role and functions of cytochromes P450.

"May the Pichia be with you"

Stellenbosch University http://scholar.sun.ac.za

APPENDIX A

DNA sequencing analyses of the human cytochrome P450 cDNAs cloned into the extracellular (pPIC9K) and intracellular (pPIC3.5K) expression vectors. Plasmid specific and gene specific primers (listed in Appendix B) were used for the different recombinant expression vectors. The transcription of the open reading frame is in bold.

Appendix A.1

DNA sequencing alignments of the pPIC9K-CYP17 expression vector harbouring the full-length CYP17 gene.

```
                   5'AOX1 primer site
pPIC9K-CYP17       TTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGA 900
Primer 5-alpha     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer STE1        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989        TTTATKAYKATATAGYTKCTTCAAATTGSNCNGGTCCAATTGCAAGTTTT 128
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                                          Start α-factor signal →
pPIC9K-CYP17       CAACTTGAGAAGATCAAAAAACAACTAATTATTCGAAGGATCCAAACGAT 950
Primer 5-alpha     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer STE1        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989        NATTTTACGCTTTNACGCCACTTARAAGATCAAAANCCACTTATTNTNGA 178
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       GAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCAT 1000
Primer 5-alpha     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer STE1        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989        AGNTCCAACGATGRATTCCNTCAATTTTKCGSAGTTTATTGSARSATCCN 228
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       TAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCG 1050
Primer 5-alpha     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer STE1        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989        CCGSATAAGGGTCCNGMNACCNTCCACKAAAATTAAKCGGCCNAANTTCC 278
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       GCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGC 1100
Primer 5-alpha     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer STE1        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989        GGTYMAAGCTGTKAYGGTTACTCAGATTNANAAGGGGATTTCGATGTTGC 328
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 1150
Primer 5-alpha     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer STE1        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989        TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 378
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                    5'α-factor primer site
pPIC9K-CYP17       CTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAA 1200
Primer 5-alpha     ~~~~~~~~~~~~~~~~~~~~~~~~~ARNNANANGGGTATCTCTCGRNAAA 25
Primer STE1        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989        CTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAA 428
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                                EcoRI ▼       → start of CYP17
pPIC9K-CYP17       AGAGAGGCTGAAGCTTACGTAGAATTCATGTGGGAGCTCGTGGCTCTCTT 1250
Primer 5-alpha     RGARRRGCTGAAGCTTACGTAGAATTCATGTGGGAGCTCGTGGCTCTCTT 75
Primer STE1        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ANANNNAAC 9
Primer 3989        AGAGAGGCTGAAGCTTACGTAGAATTCATGTGGGAGCTCGTGGCTCTCTT 478
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       GCTGCTTACCCTAGCTTATTTGTTTTGGCCCAAGAGAAGGTGCCCTGGTG 1300
Primer 5-alpha     GCTGCTTACCCTAGCTTATTTGTTTTGGCCCAAGAGAAGGTGCCCTGGTG 125
Primer STE1        ATGAATTMGCATAGCTTATTTGTTTTGGCCCAAGAGAAGGTGCCCTGGTG 59
Primer 3989        GCTGCTTACCCTAGCTTATTTGTTTTGGCCCAAGAGAAGGTGCCCTGGTG 528
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       CCAAGTACCCCAAGAGCCTCCTGTCCCTGCCCCTGGTGGGCAGCCTGCCA 1350
Primer 5-alpha     CCAAGTACCCCAAGAGCCTCCTGTCCCTGCCCCTGGTGGGCAGCCTGCCA 175
Primer STE1        CCAAGTACCCCAAGAGCCTCCTGTCCCTGCCCCTGGTGGGCAGCCTGCCA 109
Primer 3989        CCAAGTACCCCAAGAGCCTCCTGTCCCTGCCCCTGGTGGGCAGCCTGCCA 578
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       TTCCTCCCCAGACATGGCCATATGCATAACAACTTCTTCAAGCTGCAGAA 1400
Primer 5-alpha     TTCCTCCCCAGACATGGCCATATGCATAACAACTTCTTCAAGCTGCAGAA 225
Primer STE1        TTCCTCCCCAGACATGGCCATATGCATAACAACTTCTTCAAGCTGCAGAA 159
Primer 3989        TTCCTCCCCAGACATGGCCATATGCATAACAACTTCTTCAAGCTGCAGAA 628
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       AAAATATGGCCCCATCTATTCTGTTCGTATGGGCACCAAGACTACAGTGA 1450
Primer 5-alpha     AAAATATGGCCCCATCTATTCTGTTCGTATGGGCACCAAGACTACAGTGA 275
Primer STE1        AAAATATGGCCCCATCTATTCTGTTCGTATGGGCACCAAGACTACAGTGA 209
Primer 3989        AAAATATGGCCCCATCTATTCTGTTCGTATGGGCACCAAGACTACAGTGA 678
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       TTGTCGGCCACCACCAGCTGGCCAAGGAGGTGCTTATTAAGAAGGGCAAG 1500
Primer 5-alpha     TTGTCGGCCACCACCAGCTGGCCAAGGAGGTGCTTATTAAGAAGGGCAAG 325
Primer STE1        TTGTCGGCCACCACCAGCTGGCCAAGGAGGTGCTTATTAAGAAGGGCAAG 259
Primer 3989        TTGTCGGCCACCACCAGCTGGCCAAGGAGGTGCTTATTAAGAAGGGCAAG 728
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       GACTTCTCTGGGCGGCCTCAAATGGCAACTCTAGACATCGCGTCCAACAA 1550
Primer 5-alpha     GACTTCTCTGGGCGGCCTCAAATGGCAACTCTAGACATCGCGTCCAACAA 375
Primer STE1        GACTTCTCTGGGCGGCCTCAAATGGCAACTCTAGACATCGCGTCCAACAA 309
Primer 3989        GACTTCTCTGGGCGGCCTCAAATGGCAACTCTAGACATCGCGTCCAACAA 778
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       CCGTAAGGGTATCGCCTTCGCTGACTCTGGCGCACACTGGCAGCTGCATC 1600
Primer 5-alpha     CCGTAAGGGTATCGCCTTCGCTGACTCTGGCGCACACTGGCAGCTGCATC 425
Primer STE1        CCGTAAGGGTATCGCCTTCGCTGACTCTGGCGCACACTGGCAGCTGCATC 359
Primer 3989        CCGTAAGGGTATCGCCTTCGCTGACTCTGGCGCA                 812
Primer 4011        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       GAAGGCTGGCGATGGCCACCTTTGCCCTGTTCAAGGATGGCGATCAGAAG 1650
Primer 5-alpha     GAAGGCTGGCGATGGCCACCTTTGCCCTGTTCAAGGATGGCGATCAGAAG 475
Primer STE1        GAAGGCTGGCGATGGCCACCTTTGCCCTGTTCAAGGATGGCGATCAGAAG 409
Primer 3989                                                           812
Primer 4011        ~~~~~~~~GCGATGGCCACCTTTGCCCTGTTCAAGGATGGCGATCAGRAG 42
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       CTGGAGAAGATCATTTGTCAGGAAATCAGTACATTGTGTGATATGCTGGC 1700
Primer 5-alpha     CTGGAGAAGATCATTTGTCAGGAAATCAGTACATTGTGTGATATGCTGGC 525
Primer STE1        CTGGAGAAGATCATTTGTCAGGAAATCAGTACATTGTGTGATATGCTGGC 459
Primer 3989                                                           812
Primer 4011        CTGGAGAAGATCATTTGTCAGGAAATCAGTACATTGTGTGATATGCTGGC 92
Primer 8982        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ATATGTNG 8
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       CACCCACAACGGACAGTCCATAGACATCTCCTTTCCTGTCTTCGTGGCGG 1750
Primer 5-alpha     CACCCACAACGGACAGTCCATAGACATCTCCTTTCCTGTCTTCGTGGCGG 575
Primer STE1        CACCCACAACGGACAGTCCATA                             481
Primer 3989                                                           812
Primer 4011        CACCCACAACGGACAGTCCATAGACATCTCCTTTCCTGTCTTCGTGGCGG 142
Primer 8982        CCACCCACAACGGACAGCCATAGACATCTCCTTTCCTGTCTTCGTGGCGG 58
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17       TAACCAATGTCATCTCCTTGATCTGCTTCAATACCTCCTACAAGAATGGG 1800
Primer 5-alpha     TAACCAATGTCATCTCCTTGATCTGCTTCAATACCTCCTACAANAATGGG 625
Primer STE1                                                           481
Primer 3989                                                           812
Primer 4011        TAACCAATGTCATCTCCTTGATCTGCTTCAATACCTCCTACAAGAATGGG 192
Primer 8982        TAACCAATGTCATCTCCTTGATCTGCTTCAATACCTCCTACAAGAATGGG 108
Primer 1496        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
```

Stellenbosch University http://scholar.sun.ac.za

```
pPIC9K-CYP17     GACCCTGAGTTGAATGTCATACAGAATTACAATGAAGGCATCATAGACAA 1850
Primer 5-alpha   GACCCTGAGTTGAATGTCATACAGAATTACAATGAAGGCATCATAGACAN 675
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011      GACCCTGAGTTGAATGTCATACAGAATTACAATGAAGGCATCATAGACAA 242
Primer 8982      GACCCTGAGTTGAATGTCATACAGAATTACAATGAAGGCATCATAGACAA 158
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP17     CCTGAGCAAAGACAGCCTGGTGGACCTAGTCCCCTGGTTGAAGATTTTCC 1900
Primer 5-alpha   CCTGANCAAAACNGCCTGGNGGACCTAGTCCCCTGGNTGAAAATTTNCCC 725
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011      CCTGAGCAAAGACAGCCTGGTGGACCTAGTCCCCTGGTTGAAGATTTTCC 292
Primer 8982      CCTGAGCAAAGACAGCCTGGTGGACCTAGTCCCCTGGTTGAAGATTTTCC 208
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1   ~~~~~~~~~~~~~~~~YYTTNTTCNNNCCAANGGGNCNGTTTMNCAAATA 34

pPIC9K-CYP17     CCAACAAAACCCTGGAAAAATTAAAGAGCCATGTTAAAATACGAAATGAT 1950
Primer 5-alpha   CACAAACCCTGGAAAANNAAAGRCCNTGTTAAANNCNAATGATCTGGTGA 775
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011      CCAACAAAACCCTGGAAAAATTAAAGAGCCATGTTAAAATACGAAATGAT 342
Primer 8982      CCAACAAAACCCTGGAAAAATTAAAGAGCCATGTTAAAATACGAAATGAT 258
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1   AANGGCTAACCCCMNNAAACNGGGGGNCTMCCCNGNAANTTTCCCNNACC 84

pPIC9K-CYP17     CTGCTGAATAAAATACTTGAAAATTACAAGGAGAAATTCCGGAGTGACTC 2000
Primer 5-alpha   AAAAANCTTGAAATNCANGGNAATTNCGGGNKNCTTTTTNCCNATGGTGG 825
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011      CTGCTGAATAAAATACTTGAAAATTACAAGGAGAAATTCCGGAGTGACTC 392
Primer 8982      CTGCTGAATAAAATACTTGAAAATTACAAGGAGAAATTCCGGAGTGACTC 308
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1   CCNGAAAATAAANCCNMTAAAASAAAMMTTGTMAAAAAANTTAAANTACA 134

pPIC9K-CYP17     TATCACCAACATGCTGGACACACTGATGCAAGCCAAGATGAACTCAGATA 2050
Primer 5-alpha   NAMCTGGGCAGCCANWKAACTNAAWWNGGNWKGTGGCCAAAAAANYAARS 875
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011      TATCACCAACATGCTGGACACACTGATGCAAGCCAAGATGAACTCAGATA 442
Primer 8982      TATCACCAACATGCTGGACACACTGATGCAAGCCAAGATGAACTCAGATA 358
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1   GNAAANTCNGGGNNTTTTTNCCCNKATTGMGNCCCTTTAMCAAGCNAGAT 184

pPIC9K-CYP17     ATGGCAATGCTGGCCCAGATCAAGATTCAGAGCTGCTTTCAGATAACCAC 2100
Primer 5-alpha   KNTTTSANACMNTTNNCCCWNGGGMNTTNGGGYGGGGGNNMCCCCNTNGG 925
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011      ATGGCAATGCTGGCCAGATCAAGACT~CAGAGCTGCTTTCAGATA      486
Primer 8982      ATGGCAATGCTGGCCCAGATCAAGACTCAGAGCTGCTTTCAGATAACCAC 408
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1   AANTCANANNWMGCAATGGNGCCCNRATCAAAAYYNAAASNMNTTNKAAN 234

pPIC9K-CYP17     ATTCTCACCACCATAGGGGACATCTTTGGGGCTGGCGTGGAGACCACCAC 2150
Primer 5-alpha   AAAGNCCKCCTCNCNCNCCYCGGNAAKACTYCGGATRCCANGGGYN     971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982      ATTCTCACCACCATAGGGGACATCTTTGGGGCTGGCGTGGAGACCACCAC 458
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1   AACCCCATTNTCNCCNCNWWNGGGACATCTTGGGGCTGGCGTGNRRCCCC 284

pPIC9K-CYP17     CTCTGTGGTTAAATGGACCCTGGCCTTCCTGCTGCACAATCCTCAGGTGA 2200
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982      CTCTGTGGTTAAATGGACCCTGGCCTTCCTGCTGCACAATCCTCAGG    505
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1   CCCCCTCTGTGGTAAATGGACCCTGNCTTCCTGNTGCACAATCCTCAGGT 334

pPIC9K-CYP17     AGAAGAAGCTCTACGAGGAGATTGACCAGAATGTGGGTTTCAGCCGCACA 2250
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ATGTGGGTTTCAGCCGCACA 20
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX 1   AANAAAAAGCTCTACGAGGAGATGACCAGAATGTGGGTTTCAGCCGCACA 384

pPIC9K-CYP17     CCAACTATCAGTGACCGTAACCGTCTCCTCCTGCTGGAGGCCACCATCCG 2300
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      CCAACTATCAGTGACCGTAACCGTCTCCTCCTGCTGGAGGCCACCATCCG 70
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~AGGCCACCATCCG 13
Primer 3_AOX 1   CCAACTATCAGTGACGGTAACNGTCTCCTCCTGCTGNAGGCCACCATCCG 434

pPIC9K-CYP17     AGAGGTGCTTCGCCTCAGGCCCGTGGCCCCTATGCTCATCCCCCACAAGG 2350
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      AGAGGTGCTTCGCCTCAGGCCCGTGGCCCCTATGCTCATCCCCCACAAGG 120
Primer P104      AGAGGTGCTTCGCCTCAGGCCCGTGGCCCCTATGCTCATCCCCCACAAGG 63
Primer 3_AOX 1   AGAGGTGCTTCGCNTCAGGCCCGTGGCCCCTATGCTCATCCCCCACAAGG 484

pPIC9K-CYP17     CCAACGTTGACTCCAGCATCGGTGAGTTTGCTGTGGACAAGGGCACAGAA 2400
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      CCAACGTTGACTCCAGCATCGGTGAGTTTGCTGTGGACAAGGGCACAGAA 170
Primer P104      CCAACGTTGACTCCAGCATCGGTGAGTTTGCTGTGGACAAGGGCACAGAA 113
Primer 3_AOX 1   CCAACGTTGACTCCAGCATCGGTGAGTTTGCTGTGGACAAGGGCACAGAA 534

pPIC9K-CYP17     GTTATCATCAATCTGTGGGCGCTGCATCACAATGAGAAGGAGTGGCACCA 2450
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      GTTATCATCAATCTGTGGGCGCTGCATCACAATGAGAAGGAGTGGCACCA 220
Primer P104      GTTATCATCAATCTGTGGGCGCTGCATCACAATGAGAAGGAGTGGCACCA 163
Primer 3_AOX 1   GTTATCATCAATCTGTGGGCGCTGCATCACAATGAGAAGGAGTGGCACCA 584

pPIC9K-CYP17     GCCGGATCAGTTCATGCCTGAGCGTTTCTTGAATCCAGCGGGGACCCAGC 2500
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      GCCGGATCAGTTCATGCCTGAGCGTTTCTTGAATCCAGCGGGGACCCAGC 270
Primer P104      GCCGGATCAGTTCATGCCTGAGCGTTTCTTGAATCCAGCGGGGACCCAGC 213
Primer 3_AOX 1   GCCGGATCAGTTCATGCCTGAGCGTTTCTTGAATCCAGCGGGGACCCAGC 634

pPIC9K-CYP17     TCATCTCACCGTCAGTAAGCTATTTGCCCTTCGGAGCAGGACCTCGCTCC 2550
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      TCATCTCACCGTCAGTAAGCTATTTGCCCTTCGGAGCAGGACCTCGCTCC 320
Primer P104      TCATCTCACCGTCAGTAAGCTATTTGCCCTTCGGAGCAGGACCTCGCTCC 263
Primer 3_AOX 1   TCATCTCACCGTCAGTAAGCTATTTGCCCTTCGGAGCAGGACCTCGCTCC 684

pPIC9K-CYP17     TGTATAGGTGAGATCCTGGCCCGCCAGGAGCTCTTCCTCATCATGGCCTG 2600
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      TGTATAGGTGAGATCCTGGCCCGCCAGGAGCTCTTCCTCATCATGGCCTG 370
Primer P104      TGTATAGGTGAGATCCTGGCCCGCCAGGAGCTCTTCCTCATCATGGCCTG 313
Primer 3_AOX 1   TGTATAGGTGAGATCCTGGCCCGCCAGGAGCTCTTCCTCATCATGGCCTG 734

pPIC9K-CYP17     GCTGCTGCAGAGGTTCGACCTGGAGGTGCCAGATGATGGGCAGCTGCCCT 2650
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      GCTGCTGCAGAGGTTCGACCTGGAGGTGCCAGATGATGGGCAGCTGCCCT 420
Primer P104      GCTGCTGCAGAGGTTCGACCTGGAGGTGCCAGATGATGGGCAGCTGCCCT 363
Primer 3_AOX 1   GCTGCTGCAGAGGTTCGACCTGGAGGTGCCAGATGATGGGCAGCTGCCCT 784

pPIC9K-CYP17     CCCTGGAAGGCATCCCCAAGGTGGTCTTTCTGATCGACTCTTTCAAAGTG 2700
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      CCCTGGAAGGCATCCCCAAGGTGGTCTTTCTGATCGACTCTTTCAAAGTG 470
Primer P104      CCCTGGAAGGCATCCCCAAGGTGGTCTTTCTGATCGACTCTTTCAAAGTG 413
Primer 3_AOX 1   CCCTGGAAGGCATCCCCAAGGTGGTCTTTCTGATCGACTCTTTCAAAGTG 834

pPIC9K-CYP17     AAGATCAAGGTGCGCCAGGCCTGGAGGGAAGCCCAGGCTGAGGGTAGCAC 2750
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496      AAGATCAAGGTGCGCCAGGCCTGGAGGGAAACCCAAGCTGAGGGT      515
Primer P104      AAGATCAAGGTGCGCCAGGCCTMGAGGGAAGCCCAGGCTGAGGGT      458
Primer 3_AOX 1   AAGATCAAGGTGCGCCAGGCCTGGAGGGAAGCCCAGGCTGAGGGTAGCAC 884
                     6 x histidine tag
pPIC9K-CYP17     CCATCACCATCACCATCACTAAGAATTCCCTAGGGCGGCCGCGAATTAAT 2800
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496                                                         515
Primer P104                                                         458
Primer 3_AOX 1   CCATCACCATCACCATCACTAAGAATTCCCTAGGGCGGCCGCGAATTAAT 934

pPIC9K-CYP17     TCGCCTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAG 2850
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496                                                         515
Primer P104                                                         458
Primer 3_AOX 1   TCGCCTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAG 984
                                                3'AOX1 primer site
pPIC9K-CYP17     ACCGGTCTTGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCT 2900
Primer 5-alpha                                                      971
Primer STE1                                                         481
Primer 3989                                                         812
Primer 4011                                                         486
Primer 8982                                                         505
Primer 1496                                                         515
Primer P104                                                         458
Primer 3_AOX 1   ACCGMNCTTGNTAGA                                    999

pPIC9K-CYP17     GAGAGATGCAGGCT.................................... 10827
```

Stellenbosch University http://scholar.sun.ac.za

AppendixA.2

DNA sequencing alignments of the pPIC9K-truncCYP17 expression vector harbouring the truncated (Δ1-18) CYP17.

```
                  5' AOX1 primer site
pPIC9K-t.CYP17    TTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGA 900
Primer 5-alpha    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~GA 2
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                                         Start α-factor signal →
pPIC9K-t.CYP17    CAACTTGAGAAGATCAAAAAACAACTAATTATTCGAAGGATCCAAACGAT 950
Primer 5-alpha    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989       CAACTTGAGAAGATCAAAAAACAACTAATTATTCGAAGGATCCAAACGAT 52
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    GAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCAT 1000
Primer 5-alpha    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989       GAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCAT 102
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    TAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCG 1050
Primer 5-alpha    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989       TAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCG 152
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    GCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGC 1100
Primer 5-alpha    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989       GCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGC 202
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 1150
Primer 5-alpha    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3989       TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 252
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                  5' α-factor primer site
pPIC9K-t.CYP17    CTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAA 1200
Primer 5-alpha    ~~~~~~~~~~~~~~~~~~~~~~~ARNNANNANGGGTATNNTNTCGRNAAA 27
Primer 3989       CTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAA 302
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                              EcoRI ▼      → start of CYP17
pPIC9K-t.CYP17    AGAGAGGCTGAAGCTTACGTAGAATTCAAGAGAAGGTGCCCTGGTGCCAA 1250
Primer 5-alpha    RGANRRGCTGAAGCTTACGTAGAATTCAAGAGAAGGTGCCCTGGTGCCAA 77
Primer 3989       AGAGAGGCTGAAGCTTACGTAGAATTCAAGAGAAGGTGCCCTGGTGCCAA 352
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    GTACCCCAAGAGCCTCCTGTCCCTGCCCCTGGTGGGCAGCCTGCCATTCC 1300
Primer 5-alpha    GTACCCCAAGAGCCTCCTGTCCCTGCCCCTGGTGGGCAGCCTGCCATTCC 127
Primer 3989       GTACCCCAAGAGCCTCCTGTCCCTGCCCCTGGTGGGCAGCCTGCCATTCC 402
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    TCCCCAGACATGGCCATATGCATAACAACTTCTTCAAGCTGCAGAAAAAA 1350
Primer 5-alpha    TCCCCAGACATGGCCATATGCATAACAACTTCTTCAAGCTGCAGAAAAAA 177
Primer 3989       TCCCCAGACATGGCCATATGCATAACAACTTCTTCAAGCTGCAGAAAAAA 452
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    TATGGCCCCATCTATTCTGTTCGTATGGGCACCAAGACTACAGTGATTGT 1400
Primer 5-alpha    TATGGCCCCATCTATTCTGTTCGTATGGGCACCAAGACTACAGTGATTGT 227
Primer 3989       TATGGCCCCATCTATTCTGTTCGTATGGGCACCAAGACTACAGTGATTGT 502
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    CGGCCACCACCAGCTGGCCAAGGAGGTGCTTATTAAGAAGGGCAAGGACT 1450
Primer 5-alpha    CGGCCACCACCAGCTGGCCAAGGAGGTGCTTATTAAGAAGGGCAAGGACT 277
Primer 3989       CGGCCACCACCAGCTGGCCAAGGAGGTGCTTATTAAGAAGGGCAAGGACT 552
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    TCTCTGGGCGGCCTCAAATGGCAACTCTAGACATCGCGTCCAACAACCGT 1500
Primer 5-alpha    TCTCTGGGCGGCCTCAAATGGCAACTCTAGACATCGCGTCCAACAACCGT 327
Primer 3989       TCTCTGGGCGGCCTCAAATGGCAACTCTAGACATCGCGTCCAACAACCGT 602
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ACCGT 5
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    AAGGGTATCGCCTTCGCTGACTCTGGCGCACACTGGCAGCTGCATCGAAG 1550
Primer 5-alpha    AAGGGTATCGCCTTCGCTGACTCTGGCGCACACTGGCAGCTGCATCGAAG 377
Primer 3989       AAGGGTATCGCCTTCGCTGA                               622
Primer 4011       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 8982       AAGGGTATCGCCTTCGTTGACTCTGGCGCACACTGGCAGCTGCATCGAAG 55
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    GCTGGCGATGGCCACCTTTGCCCTGTTCAAGGATGGCGATCAGAAGCTGG 1600
Primer 5-alpha    GCTGGCGATGGCCACCTTTGCCCTGTTCAAGGATGGCGATCAGAAGCTGG 427
Primer 3989                                                          622
Primer 4011       ~~~~~CGATGGCCACCTTTGCCCTGTTCAAGGATGGCGATCAGAAGCTGG 45
Primer 8982       GCTGGCGATGGCCACCTTTGCCCTGTTCAAGGATGGCGATCAGAAGCTGG 105
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    AGAAGATCATTTGTCAGGAAATCAGTACATTGTGTGATATGCTGGCCACC 1650
Primer 5-alpha    AGAAGATCATTTGTCAGGAAATCAGTACATTGTGTGATATGCTGGCCACC 477
Primer 3989                                                          622
Primer 4011       AGAAGATCATTTGTCAGGAAATCAGTACATTGTGTGATATGCTGGCCACC 95
Primer 8982       AGAAGATCATTTGTCAGGAAATCAGTACATTGTGTGATATGCTGGCCACC 155
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    CACAACGGACAGTCCATAGACATCTCCTTTCCTGTCTTCGTGGCGGTAAC 1700
Primer 5-alpha    CACAACGGACAGNNYGTCCNNMSCTCCCGMKNNNNNNNNNNNNNNNNNNN 527
Primer 3989                                                          622
Primer 4011       CACAACGGACAGTCCATAGACATCTCCTTTCCTGTCTTCGTGGCGGTAAC 145
Primer 8982       CACAACGGACAGTCCATAGACATCTCCTTTCCTGTCTTCGTGGCGGTAAC 205
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    CAATGTCATCTCCTTGATCTGCTTCAATACCTCCTACAAGAATGGGGACC 1750
Primer 5-alpha    NNNNNNNNNNNNNNN                                    542
Primer 3989                                                          622
Primer 4011       CAATGTCATCTCCTTGATCTGCTTCAATACCTCCTACAAGAATGGGGACC 195
Primer 8982       CAATGTCATCTCCTTGATCTGCTTCAATACCTCCTACAAGAATGGGGACC 255
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    CTGAGTTGAATGTCATACAGAATTACAATGAAGGCATCATAGACAACCTG 1800
Primer 5-alpha                                                       542
Primer 3989                                                          622
Primer 4011       CTGAGTTGAATGTCATACAGAATTACAATGAAGGCATCATAGACAACCTG 245
Primer 8982       CTGAGTTGAATGTCATACAGAATTACAATGAAGGCATCATAGACAACCTG 305
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    AGCAAAGACAGCCTGGTGGACCTAGTCCCCTGGTTGAAGATTTTCCCCAA 1850
Primer 5-alpha                                                       542
Primer 3989                                                          622
Primer 4011       AGCAAAGACAGCCTGGTGGACCTAGTCCCCTGGTTGAAGATTTTCCCCAA 295
Primer 8982       AGCAAAGACAGCCTGGTGGACCTAGTCCCCTGGTTGAAGATTTTCCCCAA 355
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-t.CYP17    CAAAACCCTGGAAAAATTAAAGAGCCATGTTAAAATACGAAATGATCTGC 1900
Primer 5-alpha                                                       542
Primer 3989                                                          622
Primer 4011       CAAAACCCTGGAAAAATTAAAGAGCCATGTTAAAATACGAAATGATCTGC 345
Primer 8982       CAAAACCCTGGAAAAATTAAAGAGCCATGTTAAAATACGAAATGATCTGC 405
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~GGTMYSGGMTTNTTTN 16

pPIC9K-t.CYP17    TGAATAAAATACTTGAAAATTACAAGGAGAAATTCCGGAGTGACTCTATC 1950
Primer 5-alpha                                                       542
Primer 3989                                                          622
Primer 4011       TGAATAAAATACTTGAAAATTACAAGGAGAAATTCCGGAGTGACTCTATC 395
Primer 8982       TGAATAAAATACTTGAAAATTACAAGGAGAAATTCCGGAGTGACTCTATC 455
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     CCCCCACCCCGGAAAAAAAANCCNMTNAAANAAAATTTNNNNAAAAAANT 66

pPIC9K-t.CYP17    ACCAACATGCTGGACACACTGATGCAAGCCAAGATGAACTCAGATAATGG 2000
Primer 5-alpha                                                       542
Primer 3989                                                          622
Primer 4011       ACCAACATGCTGGACACACTGATGCAAGCCAAGATGAACTCAGATAATGG 445
Primer 8982       ACCAACATGCTGGACACACTGATGCAAGCCAAGATGAACTCAGATAATGG 505
Primer 1496       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3-AOX1     TAAAAAAARANAAATTNGGGGTTTTTTNCCCCNANGMGGNCCNTTTNNNN 116
```

Stellenbosch University http://scholar.sun.ac.za

```
pPIC9K-t.CYP17   CAATGCTGGCCCAGATCAAGATTCAGAGCTGCTTTCAGATAACCACATTC 2050
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011      CAATGCTGGCCCAGATCAAGACTCAGAGCTGCTTT                480
Primer 8982      CAATGCTGGCCCAGATCAAGACTCAGAGCTGCTTTCAGATAACCACATTC 555
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      ~AATGCTGGCCCAGATCAAGACTCAGAGCTGCTTTCAGATAACCACATTC 49
Primer 3-AOX1    NCNAAMNNNNYNAAAWNGGNAANGGGNCCAAAYAAAAYAAAANGNTTTTA 166

pPIC9K-t.CYP17   TCACCACCATAGGGGACATCTTTGGGGCTGGCGTGGAGACCACCACCTCT 2100
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982      TCACCACCATAGGGGACATCTTTGGGGCTGGCGTGGAGACCACCACCTCT 605
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      TCACCACCATAGGGGACATCTTTGGGGCTGGCGTGGAGACCACCACCTCT 99
Primer 3-AOX1    AAAACCCANTTTCCCCCNNAGGGGAKTTTTNGGGGGGGMGNANCCCCC 216

pPIC9K-t.CYP17   GTGGTTAAATGGACCCTGGCCTTCCTGCTGCACAATCCTCAGGTGAAGAA 2150
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982      GTGGTTAAATGGACCCTGGCCTTCCTGCTGCACAATCCTCAGGT~~~~~~ 649
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer P104      GTGGTTAAATGGACCCTGGCCTTCCTGCTGCACAATCCTCAGGTGAAGAA 149
Primer 3-AOX1    CCCNTTGGGTTAAAGNNNCNGGCCTNCNTTTMNCNAATCCNAGGNNAAAA 266

pPIC9K-t.CYP17   GAAGCTCTACGAGGAGATTGACCAGAATGTGGGTTTCAGCCGCACACCAA 2200
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 649
Primer 1496      ~~~~~~~~~~~~~~~~~~~~~~~~~~ATKTGGGTTTCAGCCGCACACCAA 24
Primer P104      GAAGCTCTACGAGGAGATTGACCAGAATGTGGGTTTCAGCCGCACACCAA 199
Primer 3-AOX1    AAAASYTTNCNAGNANNTNNCCCAAANTTGGGTTTCANCCGCNCCCCAAC 316

pPIC9K-t.CYP17   CTATCAGTGACCGTAACCGTCTCCTCCTGCTGGAGGCCACCATCCGAGAG 2250
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 649
Primer 1496      CTATCAGTGACCGTAACCGTCTCCTCCTGCTGGAGGCCACCATCCGAGAG 74
Primer P104      CTATCAGTGACCGTAACCGTCTCCTCCTGCTGGAGGCCACCATCCGAGAG 249
Primer 3-AOX1    NWTCANMGNCNGNANCCGTTTCNTCNTGTTGGAGGCCACCATCCNAAAGG 366

pPIC9K-t.CYP17   GTGCTTCGCCTCAGGCCCGTGGCCCCTATGCTCATCCCCCACAAGGCCAA 2300
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 649
Primer 1496      GTGCTTCGCCTCAGGCCCGTGGCCCCTATGCTCATCCCCCACAAGGCCAA 124
Primer P104      GTGCTTCGCCTCAGGCCCGTGGCCCCTATGCTCATCCCCCACAAGGCCAA 299
Primer 3-AOX1    TGTTTNGCCTCAGGNCCNGNGNCCCCTATGCTCATCCCCCACAAGGCCAA 416

pPIC9K-t.CYP17   CGTTGACTCCAGCATCGGTGAGTTTGCTGTGGACAAGGGCACAGAAGTTA 2350
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982      ~~~~~~~~~~~~~~~~~~~~~~                             649
Primer 1496      CGTTGACTCCAGCATCGGTGAGTTTGCTGTGGACAAGGGCACAGAAGTTA 174
Primer P104      CGTTGACTCCAGCATCGGTGAGTTTGCTGTGGACAAGGGCACAGAAGTTA 349
Primer 3-AOX1    CGTTGACTCCAGCATCGGTGAGTTTGCTGTGGACAAGGGCACAGAAGTTA 466

pPIC9K-t.CYP17   TCATCAATCTGTGGGCGCTGCATCACAATGAGAAGGAGTGGCACCAGCCG 2400
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      TCATCAATCTGTGGGCGCTGCATCACAATGAGAAGGAGTGGCACCAGCCG 224
Primer P104      TCATCAATCTGTGGGCGCTGCATCACAATGAGAAGGAGTGGCACCAGCCG 399
Primer 3-AOX1    TCATCAATCTGTGGGCGCTGCATCACAATGAGAAGGAGTGGCACCAGCCG 516

pPIC9K-t.CYP17   GATCAGTTCATGCCTGAGCGTTTCTTGAATCCAGCGGGGACCCAGCTCAT 2450
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      GATCAGTTCATGCCTGAGCGTTTCTTGAATCCAGCGGGGACCCAGCTCAT 274
Primer P104      GATCAGTTCATGCCTGAGCGTTTCTTGAATCCAGCGGGGACCCAGCTCAT 449
Primer 3-AOX1    GATCAGTTCATGCCTGAGCGTTTCTTGAATCCAGCGGGGACCCAGCTCAT 566

pPIC9K-t.CYP17   CTCACCGTCAGTAAGCTATTTGCCCTTCGGAGCAGGACCTCGCTCCTGTA 2500
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      CTCACCGTCAGTAAGCTATTTGCCCTTCGGAGCAGGACCTCGCTCCTGTA 324
Primer P104      CTCACCGTCAGTAAGCTATTTGCCCTTCGGAGCAGGACCTCGCTCCTGTA 499
Primer 3-AOX1    CTCACCGTCAGTAAGCTATTTGCCCTTCGGAGCAGGACCTCGCTCCTGTA 616

pPIC9K-t.CYP17   TAGGTGAGATCCTGGCCCGCCAGGAGCTCTTCCTCATCATGGCCTGGCTG 2550
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      TAGGTGAGATCCTGGCCCGCCAGGAGCTCTTCCTCATCATGGCCTGGCTG 374
Primer P104      TAGGTGAGATCCTGGCCCGCCAGGAGCTCTTCCTCATCATGGCCTGGCTG 549
Primer 3-AOX1    TAGGTGAGATCCTGGCCCGCCAGGAGCTCTTCCTCATCATGGCCTGGCTG 666

pPIC9K-t.CYP17   CTGCAGAGGTTCGACCTGGAGGTGCCAGATGATGGGCAGCTGCCCTCCCT 2600
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      CTGCAGAGGTTCGACCTGGAGGTGCCAGATGATGGGCAGCTGCCCTCCCT 424
Primer P104      CTGCAGAGGTTCGACCTGGAGGTGCCAGATGATGGGCAGCTGCCCTCCCT 599
Primer 3-AOX1    CTGCAGAGGTTCGACCTGGAGGTGCCAGATGATGGGCAGCTGCCCTCCCT 716

pPIC9K-t.CYP17   GGAAGGCATCCCCAAGGTGGTCTTTCTGATCGACTCTTTCAAAGTGAAGA 2650
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      GGAAGGCATCCCCAAGGTGGTCTTTCTGATCGACTCTTTCAAAGTGAAGA 474
Primer P104      GGAAGGCATCCCCAAGGTGGTCTTTCTGATCGACTCTTTCAAAGTGAAGA 649
Primer 3-AOX1    GGAAGGCATCCCCAAGGTGGTCTTTCTGATCGACTCTTTCAAAGTGAAGA 766

pPIC9K-t.CYP17   TCAAGGTGCGCCAGGCCTGGAGGGAAGCCCAGGCTGAGGGTAGCACCCAT 2700
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      TCAAGGTGCGCCAGGCCTGGAGGGAAACCCAGGCTGAGGGTAGCACCCAT 524
Primer P104      TCAAGGTGCGCCAGGCC~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 666
Primer 3-AOX1    TCAAGGTGCGCCAGGCCTGGAGGGAAGCCCAGGCTGAGGGTAGCACCCAT 816

                 6 × histidine tag
pPIC9K-t.CYP17   CACCATCACCATCACTAAGAATTCCCTAGGGCGGCCGCGAATTAATTCGC 2750
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      CACCATCACCATCACTAAGAATTCCCTAGGGCGGCCGCGAATTAATTCGC 574
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 666
Primer 3-AOX1    CACCATCACCATCACTAAGAATTCCCTAGGGCGGCCGCGAATTAATTCGC 866

pPIC9K-t.CYP17   CTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCG 2800
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      CTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCG 624
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 666
Primer 3-AOX1    CTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCG 916

                                      3'AOX1 primer site
pPIC9K-t.CYP17   GTCTTGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCTGAGA 2850
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496      GTCTTGCTAGATTCTAATCAA                              645
Primer P104      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 666
Primer 3-AOX1    MMNCTGNTASAWNNTATNGNNCA~~~~~~~~~~~~~~~~~~~~~~~~~~~ 939

pPIC9K-t.CYP17   GATGCAGGCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATA 2900
Primer 5-alpha                                                      542
Primer 3989                                                         622
Primer 4011                                                         480
Primer 8982                                                         649
Primer 1496                                                         645
Primer P104      ~                                                  666
Primer 3-AOX1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 939

pPIC9K-t.CYP17   GGATTTTTTTTGTCATTTTGTTTCTTCTCGTA.................. 10733
```

Stellenbosch University http://scholar.sun.ac.za

Appendix A. 3

DNA sequencing alignments of the pPIC9K-CYB5 expression vector harbouring the full-length CYB5.

```
pPIC9K-CYB5     CCCTGTCTTAAACCTTTTTTTTTATCATCATTATTAGCTTACTTTCATAA 850
Primer 5_AOX1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                    5' AOX1 primer site
pPIC9K-CYB5     TTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGA 900
Primer 5_AOX1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~GWNTTTMACGACTTTTAACGA 21
Primer 3_AOX1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                                           Start α-factor signal →
pPIC9K-CYB5     CAACTTGAGAAGATCAAAAAACAACTAATTATTCGAAGGATCCAAACGAT 950
Primer 5_AOX1   CARSTTGAGAAGATMAAAAAACAACTAATTATTCGAAGGATCCAAACGAT 71
Primer 3_AOX1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYB5     GAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCAT 1000
Primer 5_AOX1   GAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCAT 121
Primer 3_AOX1   ~~~~~~~~~~~~~~~~~NNNNNNNNNGAGGGGNNARGNGGNNGGGGNNGG 33

pPIC9K-CYB5     TAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCG 1050
Primer 5_AOX1   TAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCG 171
Primer 3_AOX1   NGGNGGGGGGGRANTNAACANTACAACAGAAGATAAACGGCACAAATTCC 83

pPIC9K-CYB5     GCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGC 1100
Primer 5_AOX1   GCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGC 221
Primer 3_AOX1   GGCTGAAGCTGTCATCGGTTACTCAGATTAGAAGGGGATTTCGATGNTGC 133

pPIC9K-CYB5     TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 1150
Primer 5_AOX1   TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 271
Primer 3_AOX1   TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 183

pPIC9K-CYB5     CTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAA 1200
Primer 5_AOX1   CTACTATTGCCAGCATTGCTGMTTAAGAAGAAGGGGTATCTCTCGAGAAA 321
Primer 3_AOX1   CTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAA 233

                                 EcoRI ▼      → start of CYB5
pPIC9K-CYB5     AGAGAGGCTGAAGCTTACGTAGAATTCATGGCAGAGCAGTCGGACGAGGC 1250
Primer 5_AOX1   AGAGAGGCTGAAGCTTACGTAGAATTCATGGCAGAGCAGTCGGACGAGGC 371
Primer 3_AOX1   AGAGAGGCTGAAGCTTACGTAGAATTCATGGCAGAGCAGTCGGACGAGGC 283

pPIC9K-CYB5     CGTGAAGTACTACACCCTAGAGGAGATTCAGAAGCACAACCACAGCAAGA 1300
Primer 5_AOX1   CGTGAAGTACTACACCCTAGAGGAGATTCAGAAGCACAACCACAGCAAGA 421
Primer 3_AOX1   CGTGAAGTACTACACCCTAGAGGAGATTCAGAAGCACAACCACAGCAAGA 333

pPIC9K-CYB5     GCACCTGGCTGATCCTGCACCACAAGGTGTACGATTTGACCAAATTTCTG 1350
Primer 5_AOX1   GCACCTGGCTGATCCTGCACCACAAGGTGTACGATTTGACCAAATTTCTG 471
Primer 3_AOX1   GCACCTGGCTGATCCTGCACCACAAGGTGTACGATTTGACCAAATTTCTG 383

pPIC9K-CYB5     GAAGAGCATCCTGGTGGGGAAGAAGTTTTAAGGGAACAAGCTGGAGGTGA 1400
Primer 5_AOX1   GAAGAGCATCCTGGTGGGGAAGAAGTTTTAAGGGAACAAGCTGGAGGTGA 521
Primer 3_AOX1   GAAGAGCATCCTGGTGGGGAAGAAGTTTTAAGGGAACAAGCTGGAGGTGA 433

pPIC9K-CYB5     CGCTACTGAGAACTTTGAGGATGTCGGGCACTCTACAGATGCCAGGGAAA 1450
Primer 5_AOX1   CGCTACTGAGAACTTTGAGGATGTCGGGCACTCTACAGATGCCAGGGAAA 571
Primer 3_AOX1   CGCTACTGAGAACTTTGAGGATGTCGGGCACTCTACAGATGCCAGGGAAA 483

pPIC9K-CYB5     TGTCCAAAACATTCATCATTGGGGAGCTCCATCCAGATGACAGACCAAAG 1500
Primer 5_AOX1   TGTCCAAAACATTCATCATTGGGGAGCTCCATCCAGATGACAGACCAAAG 621
Primer 3_AOX1   TGTCCAAAACATTCATCATTGGGGAGCTCCATCCAGATGACAGACCAAAG 533

pPIC9K-CYB5     TTAAACAAGCCTCCGGAAACTCTTATCACTACTATTGATTCTAGTTCCAG 1550
Primer 5_AOX1   TTAAACAAGCCTCCGGAAACTCTTATCACTACTATTGATTCTAGTTCCAG 671
Primer 3_AOX1   TTAAACAAGCCTCCGGAAACTCTTATCACTACTATTGATTCTAGTTCCAG 583

pPIC9K-CYB5     TTGGTGGACCAACTGGGTGATCCCTGCCATCTCTGCAGTGGCCGTCGCCT 1600
Primer 5_AOX1   TTGGGGGGACCAACTGGGTGATCCCTGCCATCTCTGCAGTGGCCGTCGCC 721
Primer 3_AOX1   TTGGTGGACCAACTGGGTGATCCCTGCCATCTCTGCAGTGGCCGTCGCCT 633

                                                6 × histidine tag
pPIC9K-CYB5     TGATGTATCGCCTATACATGGCAGAGGACCATCACCATCACCATCACTGA 1650
Primer 5_AOX1   TTGATGTATCGCCTATCATGGGAGAGGACCWTCACCTTCCCATNACTGAG 771
Primer 3_AOX1   TGATGTATCGCCTATACATGGCAGAGGACCATCACCATCACCATCACTGA 683

pPIC9K-CYB5     GAATTCCCTAGGGCGGCCGCGAATTAATTCGCCTTAGACATGACTGTTCC 1700
Primer 5_AOX1   ANTNCCTAGGGCGGCCGGAANTAATTCGCCTTAAMMTGACTGNTCCTNAN 821
Primer 3_AOX1   GAATTCCCTAGGGCGGCCGCGAATTAATTCGCCTTAGACATGACTGTTCC 733

pPIC9K-CYB5     TCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTAAT 1750
Primer 5_AOX1   TNNAGNNGGGCMNTNMNAAAAANCNGTCTTGNTNATTNTATNAANAGATG 871
Primer 3_AOX1   TCAGTTCAAGTTGGGCACTTACGAGAAGACCGMMNCTTGNTAGATTNCTA 783

                     3'AOX1 primer site
pPIC9K-CYB5     CAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTT 1800
Primer 5_AOX1   TMMAATGCCTTTGCCTGNAAANGCAGGCTNATTTGGANCTTTTNNTTGNA 921
Primer 3_AOX1   WNC                                                786

pPIC9K-CYB5     GATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCATTT 1850
Primer 5_AOX1   CCCWAWNGANANGATTTTTTGGMATTTGNTTNTTCGNCCNNCTGCTCC   969
Primer 3_AOX1                                                      786

pPIC9K-CYB5     TGTTTCTTCTCGTACGAGCTTGCT.......................... 9705
```

Stellenbosch University http://scholar.sun.ac.za

Appendix A.4

DNA sequencing alignments of the pPIC9K-CYP19 expression vector harbouring the full-length CYP19.

```
pPIC9K-CYP19     CCCTGTCTTAAACCTTTTTTTTTATCATCATTATTAGCTTACTTTCATAA 850
Primer 5_AOXI    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromA2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                     5' AOX1 primer site
pPIC9K-CYP19     TTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGA 900
Primer 5_AOXI    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~CGACTTTTACG 11
Primer AromA2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                                            Start α-factor signal →
pPIC9K-CYP19     CAACTTGAGAAGATCAAAAAACAACTAATTATTCGAAGGATCCAAACGAT 950
Primer 5_AOXI    ACACTTGAGAAGATCAAAAAACAACTAATTATTCGAAGGATCCAAACGAT 61
Primer AromA2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     GAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCAT 1000
Primer 5_AOXI    GAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCAT 111
Primer AromA2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     TAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCG 1050
Primer 5_AOXI    TAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCG 161
Primer AromA2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     GCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGC 1100
Primer 5_AOXI    GCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGC 211
Primer AromA2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~CGATGTTGC 9
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 1150
Primer 5_AOXI    TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 261
Primer AromA2    TGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATA 59
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     CTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAA 1200
Primer 5_AOXI    CTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAA 311
Primer AromA2    CTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAA 109
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                                            NotI ▼  start of CYP17→
pPIC9K-CYP19     AGAGAGGCTGAAGCTTACGTAGAATTCCCTAGGGCGGCCGCGGAGATGGT 1250
Primer 5_AOXI    AGAGAGGCTGAAGCTTACGTAGAATTCCCTAGGGCGGCCGCGGAGATGGT 361
Primer AromA2    AGAGAGGCTGAAGCTTACGTAGAATTCCCTAGGGCGGCCGCGGAGATGGT 159
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     TTTGGAAATGCTGAACCCGATACATTATAACATCACCAGCATCGTGCCTG 1300
Primer 5_AOXI    TTTGGAAATGCTGAACCCGATACATTATAACATCACCAGCATCGTGCCTG 411
Primer AromA2    TTTGGAAATGCTGAACCCGATACATTATAACATCACCAGCATCGTGCCTG 209
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    TTTGGAAATGCTGAACCCGATACATTATAACATCACCAGCATCGTGCCTG 50
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     AAGCCATGCCTGCTGCCACCATGCCAGTCCTGCTCCTCACTGGCCTTTTT 1350
Primer 5_AOXI    AAGCCATGCCTGCTGCCACCATGCCAGTCCTGCTCCTCACTGGCCTTTTT 461
Primer AromA2    AAGCCATGCCTGCTGCCACCATGCCAGTCCTGCTCCTCACTGGCCTTTTT 259
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    AAGCCATGCCTGCTGCCACCATGCCAGTCCTGCTCCTCACTGGCCTTTTT 100
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     CTCTTGGTGTGGAATTATGAGGGCACATCCTCAATACCAGGTCCTGGCTA 1400
Primer 5_AOXI    CTCTTGGTGTGGAATTATGAGGGCACATCCTCAATACCAGGTCCTGGCTA 511
Primer AromA2    CTCTTGGTGTGGAATTATGAGGGCACATCCTCAATACCAGGTCCTGGCTA 309
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    CTCTTGGTGTGGAATTATGAGGGCACATCCTCAATACCAGGTCCTGGCTA 150
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     CTGCATGGGAATTGGACCCCTCATCTCCCACGGCAGATTCCTGTGGATGG 1450
Primer 5_AOXI    CTGCATGGGAATTGGACCCCTCATCTCCCACGGCAGATTCCTGTGGATGG 561
Primer AromA2    CTGCATGGGAATTGGACCCCTCATCTCCCACGGCAGATTCCTGTGGATGG 359
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    CTGCATGGGAATTGGACCCCTCATCTCCCACGGCAGATTCCTGTGGATGG 200
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     GGATCGGCAGTGCCTGCAACTACTACAACCGGGTGTATGGAGAATTCATG 1500
Primer 5_AOXI    GGATCGGCAGTGCCTGCAACTACTACAACCGGGTGTATGGAGAATTCATG 611
Primer AromA2    GGATCGGCAGTGCCTGCAACTACTACAACCGGGTGTATGGAGAATTCATG 409
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    GGATCGGCAGTGCCTGCAACTACTACAACCGGGTGTATGGAGAATTCATG 250
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     CGAGTCTGGATCTCTGGAGAGGAAACACTCATTATCAGCAAGTCCTCAAG 1550
Primer 5_AOXI    CGAGTCTGGATCTCTGGAGAGGAAACACTCATTATCAGCAAGTCCTCAAG 661
Primer AromA2    CGAGTCTGGATCTCTGGAGAGGAAACACTCATTATCAGCAAGTCCTCAAG 459
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    CGAGTCTGGATCTCTGGAGAGGAAACACTCATTATCAGCAAGTCCTCAAG 300
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     TATGTTCCACATAATGAAGCACAATCATTACAGCTCTCGATTCGGCAGCA 1600
Primer 5_AOXI    TATGTCNNWNWNTDGGKRGMMRYANBTTNGNRVKKBKGNNNATNAHTAAA 711
Primer AromA2    TATGTTCCACATAATGAAGCACAATCATTACAGCTCTCGATTCGGCAGCA 509
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    TATGTTCCACATAATGAAGCACAATCATTACAGCTCTCGATTCGGCAGCA 350
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     AACTTGGGCTGCAGTGCATCGGTATGCATGAGAAAGGCATCATATTTAAC 1650
Primer 5_AOXI    TACDBANNABNYCDTDWTCNTNANNTANNSNNNNAGNNNNNNNNNNN    758
Primer AromA2    AACTTGGGCTGCAGTGCATCGGTATGCATGAGAAAGGCATCATATTTAAC 559
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    AACTTGGGCTGCAGTGCATCGGTATGCATGAGAAAGGCATCATATTTAAC 400
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     AACAATCCAGAGCTCTGGAAAACAACTCGACCCTTCTTTATGAAAGCTCT 1700
Primer 5_AOXI                                                       758
Primer AromA2    AACAATCCAGAGCTCTGGAAAACAACTCGACCCTTCTTTATGAAAGCTCT 609
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    AACAATCCAGAGCTCTGGAAAACAACTCGACCCTTCTTTATGAAAGCTCT 450
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     GTCAGGCCCCGGCCTTGTTCGTATGGTCACAGTCTGTGCTGAATCCCTCA 1750
Primer 5_AOXI                                                       758
Primer AromA2    GTCAGGCCCCGGCCTTGTTCGTATGGTCACAGTCTMGCGWCC         651
Primer AromB1    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromB2    GTCAGGCCCCGGCCTTGTTCGTATGGTCACAGTCTGTGCTGAATCCCTCA 500
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     AAACACATCTGGACAGGTTGGAGGAGGTGACCAATGAATCGGGCTATGTG 1800
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    ~~~~~ACATCTGGMCGGTTGGAGGAGGTGACCAATGAATCGGGCTATGTG 45
Primer AromB2    AAACACATCTGGACAGGTTGGAGGAGGTGACCAATGAATCGGGCTATGTG 550
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     GACGTGTTGACCCTTCTGCGTCGTGTCATGCTGGACACCTCTAACACGCT 1850
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    GACGTGTTGACCCTTCTGCGTCGTGTCATGCTGGACACCTCTAACACGCT 95
Primer AromB2    GACGTGTTGACCCTTCTGCGTCGTGTCATGCTGGACACCTCTAACACGCT 600
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     CTTCTTGAGGATCCCTTTGGACGAAAGTGCTATCGTGGTTAAAATCCAAG 1900
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    CTTCTTGAGGATCCCTTTGGACGAAAGTGCTATCGTGGTTAAAATCCAAG 145
Primer AromB2    CTTCTTGAGGATCCCTTTGGACGAAAGTGCTATCGTGGTTAAAATCCAAG 650
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     GTTATTTTGATGCATGGCAAGCTCTCCTCATCAAACCAGACATCTTCTTT 1950
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    GTTATTTTGATGCATGGCAAGCTCTCCTCATCAAACCAGACATCTTCTTT 195
Primer AromB2    GTTATTTTGATGCATGGCAAGCTCTCCTCATCAAACCAGACATCTTCTTT 700
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
```

Stellenbosch University http://scholar.sun.ac.za

```
pPIC9K-CYP19     AAGATTTCTTGGCTATACAAAAAGTATGAGAAGTCTGTCAAGGATTTGAA 2000
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    AAGATTTCTTGGCTATACAAAAAGTATGAGAAGTCTGTCAAGGATTTGAA 245
Primer AromB2    AAGATTTCTTGGCTATACAAAAAGTATGAGAAGTCTGTCAAGGATTTGAA 750
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     AGATGCCATAGAAGTTCTGATAGCAGAAAAAAGATGCAGGATTTCCACAG 2050
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    AGATGCCATAGAAGTTCTGATAGCAGAAAAAAGACGCAGGATTTCCACAG 295
Primer AromB2    AGATGCCATAGAAGTTCTGATAGCAGAAAAAAGACGCAGGATTTCCACAG 800
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     AAGAGAAACTGGAAGAATGTATGGACTTTGCCACTGAGTTGATTTTAGCA 2100
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    AAGAGAAACTGGAAGAATGTATGGACTTTGCCACTGAGTTGATTTTAGCA 345
Primer AromB2    AAGAGAAACTGGAAGAATGTATGGACTTTGCCACTGAGTTGATTTAGCAG 850
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     GAGAAACGTGGTGACCTGACAAGAGAGAATGTGAACCAGTGCATATTGGA 2150
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    GAGAAACGTGGTGACCTGACAAGAGAGAATGTGAACCAGTGCATATTGGA 395
Primer AromB2    AGAAACGGAC                                         860
Primer AromaC1   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~CATATTGGA 9
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC9K-CYP19     AATGCTGATCGCAGCTCCTGACACCATGTCTGTCTCTTTGTTCTTCATGC 2200
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    AATGCTGATCGCAGCTCCTGACACCATGTCTGTCTCTTTGTTCTTCATGC 445
Primer AromB2                                                       860
Primer AromaC1   AATGCTGAYCGCAGCTCCTGAMACCATGTCTGTCTCTTTGTTCTTCATGC 59
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     ~~~~~~~~~~~~~NWTTGNAAMSMRTNSAGNTCTGCCCCAAGTTGTTTTT 37

pPIC9K-CYP19     TATTTCTCATTGCAAAGCACCCTAATGTTGAAGAGGCAATAATAAAGGAA 2250
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    TATTTCTCATTGCAAAGCACCCTAATGTTGAAGAGGCAATAATAAAGGAA 495
Primer AromB2                                                       860
Primer AromaC1   TATTTCTCATTGCAAAGCACCCTAATGTTGAAGAGGCAATAATAAAGGAA 109
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     GTTTTCANGNATTTNCATGCAAANCCCCTANTNNNNANRGGCAATATAAN 87

pPIC9K-CYP19     ATCCAGACTGTTATTGGTGAGAGAGACATAAAGATTGATGATATACAAAA 2300
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    ATCCAGACTGTTATTGGTGAGAGAGACATAAAGATTGATGATATACAAAA 545
Primer AromB2                                                       860
Primer AromaC1   ATCCAGACTGTTATTGGTGAGAGAGACATAAAGATTGATGATATACAAAA 159
Primer AromaC2   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
Primer 3_AOX     GNAATCCAGNCTGTTTTGGTGAGGGNGCKWAAAGNTTGATGATWTCNAAA 137

pPIC9K-CYP19     ATTAAAAGTGATGGAAAACTTCATTTATGAGAGCATGCGGTACCAGCCTG 2350
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    ATTAAAAGTGATGGAAAACTTCATTTATGAGAGCATGCGGTACCAGCCTG 595
Primer AromB2                                                       860
Primer AromaC1   ATTAAAAGTGATGGAAAACTTCATTTATGAGAGCATGCGGTACCAGCCTG 209
Primer AromaC2   ~~~~~~~~~~~~NCNNNGSGTTHGCTATGAGAGCATGCGGTACCAGCCTG 38
Primer 3_AOX     ATTAAAAGTGATGGAAAACTTCATTTATGAGAGCATGCGGTACCAGCCTG 187

pPIC9K-CYP19     TCGTGGACTTGGTCATGCGCAAAGCCTTAGAAGATGATGTAATCGATGGC 2400
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    TCGTGGACTTGGTCATGCGCAAAGCCTTAGAAGATGATGTAATCGATGGC 645
Primer AromB2                                                       860
Primer AromaC1   TCGTGGACTTGGTCATGCGCAAAGCCTTAGAAGATGATGTAATCGATGGC 259
Primer AromaC2   TCGTGGACTTGGTCATGCGCAAAGCCTTAGAAGATGATGTAATCGATGGC 88
Primer 3_AOX     TCGTGGACTTGGTCATGCGCAAAGCCTTAGAAGATGATGTAATCGATGGC 237

pPIC9K-CYP19     TACCCAGTGAAAAAGGGGACAAACATTATCCTGAATATTGGAAGGATGCA 2450
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    TACCCAGTGNAAAAGGGGACAAACATTATCCTGAATATTGGAAGGATGCA 695
Primer AromB2                                                       860
Primer AromaC1   TACCCAGTGAAAAAGGGGACAAACATTATCCTGAATATTGGAAGGATGCA 309
Primer AromaC2   TACCCAGTGAAAAAGGGGACAAACATTATCCTGAATATTGGAAGGATGCA 138
Primer 3_AOX     TACCCAGTGAAAAAGGGGACAAACATTATCCTGAATATTGGAAGGATGCA 287

pPIC9K-CYP19     CAGACTCGAGTTTTTCCCCAAACCCAATGAATTTACTCTTGAAAATTTTG 2500
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    CAGACTCGAGTTTTTCCCCAACCCAATGANTTACTCTTGAAANTTTGCAA 745
Primer AromB2                                                       860
Primer AromaC1   CAGACTCGAGTTTTTCCCCAAACCCAATGAATTTACTCTTGAAAATTTTG 359
Primer AromaC2   CAGACTCGAGTTTTTCCCCAAACCCAATGAATTTACTCTTGAAAATTTTG 188
Primer 3_AOX     CAGACTCGAGTTTTTCCCCAAACCCAATGAATTTACTCTTGAAAATTTTG 337

pPIC9K-CYP19     CAAAGAATGTTCCTTATAGGTACTTTCAGCCATTTGGCTTTGGGCCCCGT 2550
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    AAAAGTGCYCNTANGNGNACTTNMCCCATTTGGNTTNGGNCCCGTGGCTG 795
Primer AromB2                                                       860
Primer AromaC1   CAAAGAATGTTCCTTATAGGTACTTTCAGCCATTTGGCTNKGGNGNGNN  408
Primer AromaC2   CAAAGAATGTTCCTTATAGGTACTTTCAGCCATTTGGCTTTGGGCCCCGT 238
Primer 3_AOX     CAAAGAATGTTCCTTATAGGTACTTTCAGCCATTTGGCTTTGGGCCCCGT 387

pPIC9K-CYP19     GGCTGTGCAGGAAAGTACATCGCCATGGTGATGATGAAAGCCATCCTCGT 2600
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    TGCANGAAAGTCATCGCCATGGTGATGATGAAGCATCCTCGTNMMCTTCT 845
Primer AromB2                                                       860
Primer AromaC1                                                      408
Primer AromaC2   GGCTGTGCAGGAAAGTACATCGCCATGGTGATGATGAAAGCCATCCTCGT 288
Primer 3_AOX     GGCTGTGCAGGAAAGTACATCGCCATGGTGATGATGAAAGCCATCCTCGT 437

pPIC9K-CYP19     TACACTTCTGAGACGATTCCACGTGAAGACATTGCAAGGACAGTGTGTTG 2650
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    GAAACGATCCACGTGAAACNTTGCANGACRGGTGTTGANNSAWACAAAAA 895
Primer AromB2                                                       860
Primer AromaC1                                                      408
Primer AromaC2   TACACTTCTGAGACGATTCCACGTGAAGACATTGCAAGGACAGTGTGTTG 338
Primer 3_AOX     TACACTTCTGAGACGATTCCACGTGAAGACATTGCAAGGACAGTGTGTTG 487

pPIC9K-CYP19     AGAGCATACAGAAGATACACGACTTGTCCTTGCACCCAGATGAGACTAAA 2700
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    NACCGACTNNNCTTGCCCNAATGAACTAAAACATGYTGNAATRYCTTNCC 945
Primer AromB2                                                       860
Primer AromaC1                                                      408
Primer AromaC2   AGAGCATACAGAAGATACACGACTTGTCCTTGCACCCAGATGAGACTAAA 388
Primer 3_AOX     AGAGCATACAGAAGATACACGACTTGTCCTTGCACCCAGATGAGACTAAA 537

pPIC9K-CYP19     AACATGCTGGAAATGATCTTTACCCCAAGAAACTCAGACAGGTGTCTGGA 2750
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    CCANAAYYCNANGGGGTYGGANCCCTCCCCCCNTTACCGGGGAATCCNG  995
Primer AromB2                                                       860
Primer AromaC1                                                      408
Primer AromaC2   AACATGCTGAAA                                       400
Primer 3_AOX     AACATGCTGGAAATGATCTTTACCCCAAGAAACTCAGACAGGTGTCTGGA 587

                 6 x histidine tag
pPIC9K-CYP19     ACACCATCACCATCACCATTAACCGCGGGAATTCGCGGCCGCGAATTAAT 2800
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    CCGNAATANTCCCTNAAATGAYKKTNNCANTNAGTGGGMYNCANAAANGG 1045
Primer AromB2                                                       860
Primer AromaC1                                                      408
Primer AromaC2                                                      400
Primer 3_AOX     ACACCATCACCATCACCATTAACCGCGGGAATTCGCGGCCGCGAATTAAT 637

pPIC9K-CYP19     TCGCCTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAG 2850
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    NNTSYAATYTAACAAGRGKCAAAGCCTTNCCGAAGGMGGTYYNTTTNACT 1095
Primer AromB2                                                       860
Primer AromaC1                                                      408
Primer AromaC2                                                      400
Primer 3_AOX     TCGCCTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAG 687

                                           3'AOX1 primer site
pPIC9K-CYP19     ACCGGTCTTGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCT 2900
Primer 5_AOXI                                                       758
Primer AromA2                                                       651
Primer AromB1    TTTTNRANNAN                                        1106
Primer AromB2                                                       860
Primer AromaC1                                                      408
Primer AromaC2                                                      400
Primer 3_AOX     ACCGMCTTGCTAG~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 700

pPIC9K-CYP19     ACACTGAAAAATAACAGTTATTATTCG .......................10827
```

Stellenbosch University http://scholar.sun.ac.za

Appendix A.5

DNA sequencing alignments of the pPIC3.5K-CYP17 expression vector harbouring the full-length CYP17.

```
                       5' AOX1 primer site
pPIC3.5K-CYP17   TTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGA 900
5_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989             TTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGA 334
8982             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                                                           EcoRI ▼
pPIC3.5K-CYP17   CAACTTGAGAAGATCAAAAAACAACTAATTATTCGAAGGATCCTACGTAG 950
5_AOX1           ~~~~~~~~~~~~~~~~AAAAACAACTAATTATTCGAAGGATCCTACGTAG 34
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989             CAACTTGAGAAGATCAAAAAACAACTAATTATTCGAAGGATCCTACGTAG 384
8982             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

                       → start of CYP17
pPIC3.5K-CYP17   AATTCATGTGGGAGCTCGTGGCTCTCTTGCTGCTTACCCTAGCTTATTTG 1000
5_AOX1           AATTCATGTGGGAGCTCGTGGCTCTCTTGCTGCTTACCCTAGCTTATTTG 84
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989             AATTCATGTGGGAGCTCGTGGCTCTCTTGCTGCTTACCCTAGCTTATTTG 434
8982             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   TTTTGGCCCAAGAGAAGGTGCCCTGGTGCCAAGTACCCCAAGAGCCTCCT 1050
5_AOX1           TTTTGGCCCAAGAGAAGGTGCCCTGGTGCCAAGTACCCCAAGAGCCTCCT 134
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989             TTTTGGCCCAAGAGAAGGTGCCCTGGTGCCAAGTACCCCAAGAGCCTCCT 484
8982             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   GTCCCTGCCCCTGGTGGGCAGCCTGCCATTCCTCCCCAGACATGGCCATA 1100
5_AOX1           GTCCCTGCCCCTGGTGGGCAGCCTGCCATTCCTCCCCAGACATGGCCATA 184
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989             GTCCCTGCCCCTGGTGGGCAGCCTGCCATTCCTCCCCAGACATGGCCATA 534
8982             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   TGCATAACAACTTCTTCAAGCTGCAGAAAAAATATGGCCCCATCTATTCT 1150
5_AOX1           TGCATAACAACTTCTTCAAGCTGCAGAAAAAATATGGCCCCATCTATTCT 234
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989             TGCATAACAACTTCTTCAAGCTGCAGAAAAAATATGGCCCCATCTATTCT 584
8982             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   GTTCGTATGGGCACCAAGACTACAGTGATTGTCGGCCACCACCAGCTGGC 1200
5_AOX1           GTTCGTATGGGCACCAAGACTACAGTGATTGTCGGCCACCACCAGCTGGC 284
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989             GTTCGTATGGGCACCAAGACTACAGTGATTGTCGGCCACCACCAGCTGGC 634
8982             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   CAAGGAGGTGCTTATTAAGAAGGGCAAGGACTTCTCTGGGCGGCCTCAAA 1250
5_AOX1           CAAGGAGGTGCTTATTAAGAAGGGCAAGGACTTCTCTGGGCGGCCTCAAA 334
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989             CAAGGAGGTGCTTATTAAGAAGGGCAAGGACTTCTCTGGGCGGCCTCAAA 684
8982             ~~~~~~~~~~~~GGTGCTTATTAAGAAGGGCAAGGACTTCTCTGGGCGGC 38
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   TGGCAACTCTAGACATCGCGTCCAACAACCGTAAGGGTATCGCCTTCGCT 1300
5_AOX1           TGGCAACTCTAGACATCGCGTCCAACAACCGTAAGGGTATCGCCTTCGCT 384
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989             TGGCAA                                             690
8982             CTCAAATGGCAACTCTAGACATCGCGTCCAACAACCGTAAGGGTATCGCC 88
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   GACTCTGGCGCACACTGGCAGCTGCATCGAAGGCTGGCGATGGCCACCTT 1350
5_AOX1           GACTCTGGCGCACACTGGCAGCTGCATCGAAGGCTGGCGATGGCCACCTT 434
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
3989                                                                690
8982             TTCGCTGACTCTGGCGCACACTGGCAGCTGCATCGAAGGCTGGCGATGGC 138
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   TGCCCTGTTCAAGGATGGCGATCAGAAGCTGGAGAAGATCATTTGTCAGG 1400
5_AOX1           TGCCCTGTTCAARGATGGCGATCAGAAGCTGGAGAAGATCATTTGTCAGG 484
4011             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~CATTTGTCAGG 11
3989                                                                690
8982             CACCTTTGCCCTGTTCAAGGATGGCGATCAGAAGCTGGAGAAGATCATTT 188
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   AAATCAGTACATTGTGTGATATGCTGGCCACCCACAACGGACAGTCCATA 1450
5_AOX1           AAATCAGTACATTGTGTGATATGCTGGCCACCCACAACGGACAGTCCATA 534
4011             AAATCAGTACATTGTGTGATATGCTGGCCACCCACAACGGACAGTCCATA 61
3989                                                                690
8982             GTCAGGAAATCAGTACATTGTGTGATATGCTGGCCACCCACAACGGACAG 238
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   GACATCTCCTTTCCTGTCTTCGTGGCGGTAACCAATGTCATCTCCTTGAT 1500
5_AOX1           GACATCTCCTTTCCTGTCTTCGTGGCGGTAACCAATGTCATCTCCTTGAT 584
4011             GACATCTCCTTTCCTGTCTTCGTGGCGGTAACCAATGTCATCTCCTTGAT 111
3989                                                                690
8982             TCCATAGACATCTCCTTTCCTGTCTTCGTGGCGGTAACCAATGTCATCTC 288
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   CTGCTTCAATACCTCCTACAAGAATGGGGACCCTGAGTTGAATGTCATAC 1550
5_AOX1           CTGCTTCANTACCTCCTACAAGAATGGGGACCCTGAGTTGAATGTCATAC 634
4011             CTGCTTCAATACCTCCTACAAGAATGGGGACCCTGAGTTGAATGTCATAC 161
3989                                                                690
8982             CTTGATCTGCTTCAATACCTCCTACAAGAATGGGGACCCTGAGTTGAATG 338
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   AGAATTACAATGAAGGCATCATAGACAACCTGAGCAAAGACAGCCTGGTG 1600
5_AOX1           AGAATTACANTGAAGGCATCATAGACAACCTGAGCAAAGACAGCCTGGTG 684
4011             AGAATTACAATGAAGGCATCATAGACAACCTGAGCAAAGACAGCCTGGTG 211
3989                                                                690
8982             TCATACAGAATTACAATGAAGGCATCATAGACAACCTGAGCAAAGACAGC 388
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   GACCTAGTCCCCTGGTTGAAGATTTTCCCCAACAAAACCCTGGAAAAATT 1650
5_AOX1           GACCTAGTCCCCTGGTTGAAGATTTTCCC                      713
4011             GACCTAGTCCCCTGGTTGAAGATTTTCCCCAACAAAACCCTGGAAAAATT 261
3989                                                                690
8982             CTGGTGGACCTAGTCCCCTGGTTGAAGATTTTCCCCAACAAAACCCTGGA 438
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   AAAGAGCCATGTTAAAATACGAAATGATCTGCTGAATAAAATACTTGAAA 1700
5_AOX1                                                              713
4011             AAAGAGCCATGTTAAAATACGAAATGATCTGCTGAATAAAATACTTGAAA 311
3989                                                                690
8982             AAAATTAAAGAGCCATGTTAAAATACGAAATGATCTGCTGAATAAAATAC 488
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   ATTACAAGGAGAAATTCCGGAGTGACTCTATCACCAACATGCTGGACACA 1750
5_AOX1                                                              713
4011             ATTACAAGGAGAAATTCCGGAGTGACTCTATCACCAACATGCTGGACACA 361
3989                                                                690
8982             TTGAAAATTACAAGGAGAAATTCCGGAGTGACTCTATCACCAACATGCTG 538
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   CTGATGCAAGCCAAGATGAACTCAGATAATGGCAATGCTGGCCCAGATCA 1800
5_AOX1                                                              713
4011             CTGATGCAAGCCAAGATGAACTCAGATAATGGCAATGCTGGCCCAGATCA 411
3989                                                                690
8982             GACACACTGATGCAAGCCAAGATGAACTCAGATAATGGCAATGCTGGCCC 588
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   AGATTCAGAGCTGCTTTCAGATAACCACATTCTCACCACCATAGGGGACA 1850
5_AOX1                                                              713
4011             AGACTCAGAGCTGCTTTCAGATAACCACATTCTCACCACCATAGGGGACA 461
3989                                                                690
8982             AGATCAAGACTCAGAGCTGCTTTCAGATAACCACATTCTCACCACCATAG 638
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   TCTTTGGGGCTGGCGTGGAGACCACCACCTCTGTGGTTAAATGGACCCTG 1900
5_AOX1                                                              713
4011             TCTTTGGGGCTGGCGTGGAGACCACCACCTCTGTGGTTAAATGGACCCTG 511
3989                                                                690
8982             GGGACATCTTTGGGGCTGGCGTGGAGACCACCAC                 672
3_AOX1           ~~~~~~~~~~~~~~~~~~~~~~~~~~~CCTCTGTGGTTAAATGGACCCTG 23
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   GCCTTCCTGCTGCACAATCCTCAGGTGAAGAAGAAGCTCTACGAGGAGAT 1950
5_AOX1                                                              713
4011             GCCTTCCTGCTGCACAATCCTCAGGTGAAGAANAAGCTCTACGAGGAGAT 561
3989                                                                690
8982                                                                672
3_AOX1           GCCTTCCTGCTGCACAATCCTCAGGTGAAGAAGAAGCTCTACGAGGAGAT 73
1496             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ 1

pPIC3.5K-CYP17   TGACCAGAATGTGGGTTTCAGCCGCACACCAACTATCAGTGACCGTAACC 2000
5_AOX1                                                              713
4011             TGACCAGAATGTGGGTTTCAGCCGCACACCAACTATCAGTGACCGTAACC 611
3989                                                                690
8982                                                                672
3_AOX1           TGACCAGAATGTGGGTTTCAGCCGCACACCAACTATCAGTGACCGTAACC 123
1496             ~~~~~~~~~AGTGGGTTTCAGCCGCACACCAACTATCAGTGACCGTAACC 41

pPIC3.5K-CYP17   GTCTCCTCCTGCTGGAGGCCACCATCCGAGAGGTGCTTCGCCTCAGGCCC 2050
5_AOX1                                                              713
4011             GTCT                                               615
3989                                                                690
8982                                                                672
3_AOX1           GTCTCCTCCTGCTGGAGGCCACCATCCGAGAGGTGCTTCGCCTCAGGCCC 173
1496             GTCTCCTCCTGCTGGAGGCCACCATCCGAGAGGTGCTTCGCCTCAGGCCC 91

pPIC3.5K-CYP17   GTGGCCCCTATGCTCATCCCCCACAAGGCCAACGTTGACTCCAGCATCGG 2100
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           GTGGCCCCTATGCTCATCCCCCACAAGGCCAACGTTGACTCCAGCATCGG 223
1496             GTGGCCCCTATGCTCATCCCCCACAAGGCCAACGTTGACTCCAGCATCGG 141

pPIC3.5K-CYP17   TGAGTTTGCTGTGGACAAGGGCACAGAAGTTATCATCAATCTGTGGGCGC 2150
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           TGAGTTTGCTGTGGACAAGGGCACAGAAGTTATCATCAATCTGTGGGCGC 273
1496             TGAGTTTGCTGTGGACAAGGGCACAGAAGTTATCATCAATCTGTGGGCGC 191

pPIC3.5K-CYP17   TGCATCACAATGAGAAGGAGTGGCACCAGCCGGATCAGTTCATGCCTGAG 2200
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           TGCATCACAATGAGAAGGAGTGGCACCAGCCGGATCAGTTCATGCCTGAG 323
1496             TGCATCACAATGAGAAGGAGTGGCACCAGCCGGATCAGTTCATGCCTGAG 241

pPIC3.5K-CYP17   CGTTTCTTGAATCCAGCGGGGACCCAGCTCATCTCACCGTCAGTAAGCTA 2250
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           CGTTTCTTGAATCCAGCGGGGACCCAGCTAATCTCACCGTCAGTAAGCTA 373
1496             CGTTTCTTGAATCCAGCGGGGACCCAGCTAATCTCACCGTCAGTAAGCTA 291
```

Stellenbosch University http://scholar.sun.ac.za

```
pPIC3.5K-CYP17   TTTGCCCTTCGGAGCAGGACCTCGCTCCTGTATAGGTGAGATCCTGGCCC 2300
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           TTTGCCCTTCGGAGCAGGACCTCGCTCCTGTATAGGTGAGATCCTGGCCC 423
1496             TTTGCCCTTCGGAGCAGGACCTCGCTCCTGTATAGGTGAGATCCTGGCCC 341

pPIC3.5K-CYP17   GCCAGGAGCTCTTCCTCATCATGGCCTGGCTGCTGCAGAGGTTCGACCTG 2350
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           GCCAGGAGCTCTTCCTCATCATGGCCTGGCTGCTGCAGAGGTTCGACCTG 473
1496             GCCAGGAGCTCTTCCTCATCATGGCCTGGCTGCTGCAGAGGTTCGACCTG 391

pPIC3.5K-CYP17   GAGGTGCCAGATGATGGGCAGCTGCCCTCCCTGGAAGGCATCCCCAAGGT 2400
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           GAGGTGCCAGATGATGGGCAGCTGCCCTCCCTGGAAGGCATCCCCAAGGT 523
1496             GAGGTGCCAGATGATGGGCAGCTGCCCTCCCTGGAAGGCATCCCCAAGGT 441

pPIC3.5K-CYP17   GGTCTTTCTGATCGACTCTTTCAAAGTGAAGATCAAGGTGCGCCAGGCCT 2450
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           GGTCTTTCTGATCGACTCTTTCAAAGTGAAGATCAAGGTGCGCCAGGCCT 573
1496             GGTCTTTCTGATCGACTCTTTCAAAGTGAAGATCAAGGTGCGCCAGGCCT 491

                                                6 × histidine tag
pPIC3.5K-CYP17   GGAGGGAAGCCCAGGCTGAGGGTAGCACCCATCACCATCACCATCACTAA 2500
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           GGAGGGAAGCCCAGGCTGAGGGTAGCACCCATCACCATCACCATCACTAA 623
1496             GGAGGGAAGCCCAGGCTGAGGGTAGCACCCATCACCATCACCATCACTAA 541

pPIC3.5K-CYP17   GAATTCCCTAGGGCGGCCGCGAATTAATTCGCCTTAGACATGACTGTTCC 2550
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           GAATTCCCTAGGGCGGCCGCGAATTAATTCGCCTTAGACATGACTGTTCC 673
1496             GAATTCCCTAGGGCGGCCGCGAATTAATTCGCCTTAGACATGACTGTTCC 591

pPIC3.5K-CYP17   TCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTAAT 2600
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1           TCAGTTCAAGTTGGGCACTTACGAGAAGACCGGT                 707
1496             TCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTATT 641

pPIC3.5K-CYP17   CAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTT 2650
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1                                                              707
1496             CAAGAGG                                            648

pPIC3.5K-CYP17   GATACTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCATTT 2700
5_AOX1                                                              713
4011                                                                615
3989                                                                690
8982                                                                672
3_AOX1                                                              707
1496                                                                648

pPIC3.5K-CYP17   ATTCGACCGCC                                        10561
```

Stellenbosch University http://scholar.sun.ac.za

APPENDIX B

Primers used in this study.

Primer	Gene specificity	Primer sequence	Type of Primer
STE1		5'-ATGCGAATTCATGTGGGAGCTCGTGGCTCTC-3'	Sense
4011		5'-TGGCAGCTCATCGAAGGCT-3'	Sense
1496		5'-GAAGC TCTACGAGGAGATTGACCAG-3'	Sense
574	CYP17	5'-TAGAAGCTTATGGTGATGGTGATGGGTGCTACCCTCAGCCTG-3'	Antisense
3989		5'-AGCCTTCGATGCAGCTGCCAG-3'	Antisense
8982		5'-CTGGTCAATCTCCTCGTAGAGCT TC-3'	Antisense
STE2		5'-AGCTAAGCTTTTAGGTGCTACCCTCAGCCTG-3'	Antisense
YαN-21	pFLAG-1	5'-AGCACAAATAACGGGTTATTG-3'	Sense
YαC-21		5-TACAGACGCGTGTACGCATGT-3'	Antisense
P104	CYP17	5'-CCGGAATTCTTATGATGGTGATGGTGATGGGTGCT-3'	Antisense
α-factorP	*MFα-1*	5'-TACTATTGCCAGCATTGCTGC-3'	Sense
3'AOX1P	3'-*AOX1*	5'-GCAAATGGCATTCTGACATCC-3'	Antisense
P103	TruncCYP17	5'-ATGCGAATTCAAGAGAAGGTGCCCTGGTGC-3'	Sense
P107		5'-AGGCGAATTCATGGCAGAGCAGTCGGAC-3'	Sense
P108	CYB5	5'-GTGATGGTGATGGTGATGGTCCTCTGCCATGTATAGG-3'	Antisense
P109		5'-AGGCGAATTCATGGCAGA-3'	Sense
P110		5'-CCGGAATTCTCAGTGATGGTGATGGTGATGGT-3'	Antisense
P19A1		5'-GGTCAAGGAACACAAGATGG-3'	Sense
P19A2		5'-ATGGTGATGGTGATGGTGTTCCAGACACCTTCTGA-3'	Antisense
P19B1		5'-TCCCCGCGGAGATGGTTTTGGAAATGCTGAA-3'	Sense
P19B2		5'-TCCCCGCGGTTAATGGTGATGGTGATGGTGTT-3'	Antisense
AromB1	CYP19	5'-GTCACAGTCTGTGCTGAATC-3'	Sense
AromC1		5'-GGTGACCTGACAAGAGAGAA-3'	Sense
AromA2		5'-CTCCAACCTGTCCAGATGTG-3'	Antisense
AromB2		5'-GCACTGGTTCACATTCTCTC-3'	Antisense
AromC2		5'-CACCTGTCTGAGTTTCTTGG-3'	Antisense
P105	CYP17	5'-ATGCGAATTCATGTGGGAGCTCGTG-3'	Sense
P106		5'-GTACGCGGCCGCGAATTCTTAGTGATGGTGAT-3'	Antisense
5'AOX1P	5'-*AOX1*	5'-GACTGGTTCCAATTGACAAGC-3'	Sense