HIERDIE EKSEMPLAAR MAG ONDER GEEEN OMSTANDIGHEDDE UIT DIE BIBLIOTEEK VERWYDER WORD NIE.
Differential transcription of \textit{CYP52} genes of \textit{Yarrowia lipolytica} during growth on hydrocarbons

by

Michael Bertram Drennan

Submitted in fulfilment of the requirements for the degree

\textbf{MAGISTER SCIENTIA}

in the

Department of Microbiology and Biochemistry
Faculty of Natural Sciences
University of the Free State
Bloemfontein
Republic of South Africa

November, 2000

Study leader : Prof. M.S. Smit
Co-study leader : Prof. H.-G. Patterton
Acknowledgements

“A well-known scientist (some say it was Bertrand Russell) once gave a public lecture on astronomy. He described how the earth orbits around the sun and how the sun, in turn, orbits around the centre of a vast collection of stars called our galaxy. At the end of the lecture, a little old lady at the back of the room got up and said: “What you have told us is rubbish. The world is really a flat plate supported on the back of a giant tortoise.” The scientist gave a superior smile before replying, “What is the tortoise standing on?” “You’re very clever, young man, very clever,” said the old lady. “But it’s turtles all the way down!”

Stephen W. Hawking

I am very sure that all scientists endeavor to be pioneers, albeit a long and hard road! This study has certainly taught me patience, of which I initially thought I had a lot of! I am indebted to my study leaders, Prof. M.S. Smit and Prof. H.-G. Patterton, who not only assisted me at the worst of times, but who probably had their patience tried more so than I had mine! I would like to thank my father, mother, brother, sister and Lieschen from the bottom of my heart for having all the faith in the world in me, without which, I would certainly have fallen by the wayside. A special word of thanks goes to my fellow colleagues who were always more than willing to help in any way they could. I also wish to thank Sasol Technology Pty Ltd for financial support.

Finally, a word for future pioneers:

The Road goes ever on and on
Down from the door where it began
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with weary feet,
Until it joins some larger way,
Where many paths and errands meet.
And whither then? I cannot say.

John R.R. Tolkien
# Table of Contents

**Acknowledgements** 2

**Table of Contents** 3

**List of Abbreviations** 6

**List of Figures** 8

**List of Tables** 10

**Chapter One** 11

1.1 General Introduction 11

1.2 Aims of the research project 12

**Chapter Two: The regulation of Cytochrome P450 monooxygenases in prokaryotic and eukaryotic organisms – a literature overview** 13

2.1 Introduction 13

2.2 The P450 gene superfamily 15

2.2.1 Summary of the P450 gene nomenclature 15

2.2.2 Regulation of the P450 gene superfamily 16

2.3 Bacterial Cytochrome P450 monooxygenases 16

2.3.1 The *Pseudomonas putida* Cytochrome P450<sub>cam</sub> 17

2.3.2 The *Bacillus megaterium* Cytochrome P450<sub>BM-3</sub> 18

2.3.3 Other bacterial Cytochrome P450 monooxygenases 19

2.4 Mammalian Cytochrome P450 monooxygenases 20

2.4.1 Aromatic-hydrocarbon-inducible Cytochrome P450 monooxygenases 20

2.4.1.1 The P450I gene family 20

2.4.1.2 Regulation of the *CYP1* gene family 21

2.4.2 Phenobarbital-inducible Cytochrome P450 monooxygenases 23

2.4.2.1 The P450II gene family 23

2.4.2.1.1 The *CYP2A* gene subfamily 23

2.4.2.1.2 The *CYP2B* gene subfamily 24

2.4.2.1.3 The *CYP2C* gene subfamily 24

2.4.2.1.4 The *CYP2D* gene subfamily 25

2.4.2.1.5 The *CYP2E* gene subfamily 25

2.4.2.2 Regulation of the *CYP2* gene family 26

2.4.3 Glucocorticoid-inducible Cytochrome P450 monooxygenases 29

2.4.3.1 The P450III gene family 29
2.4.3.2 Regulation of the CYP3 gene family
2.4.4 Peroxisome Proliferator-inducible Cytochrome P450 monooxygenases
  2.4.4.1 The P450IV gene family
  2.4.4.2 Regulation of the CYP4 gene family

2.5 Fungal Cytochrome P450 monooxygenases
  2.5.1 The CYP51 gene family involved in ergosterol biosynthesis
  2.5.2 The CYP52 gene family involved in alkane utilization

2.6 Conclusions

Chapter Three: Materials and Methods

3.1 Materials
  3.1.1 Strains, media and growth conditions
  3.1.2 Chemicals, enzymes and kits

3.2 Methods
  3.2.1 Isolation of chromosomal DNA
  3.2.2 Isolation of total RNA
  3.2.3 Primers for amplifying CYP52 and Actin gene fragments from
      Y. lipolytica UOFS Y-0164
  3.2.4 PCR amplification of CYP52 gene fragments from Y. lipolytica UOFS
      Y-0164 using ALK-1/ALK-REV to ALK-8/ALK-REV primer pairs
  3.2.5 PCR amplification of Actin gene fragments from Y. lipolytica UOFS
      Y-0164 using ACT(F)/ACT(R) primer pair
  3.2.6 Cloning and DNA sequencing
  3.2.7 RT-PCR amplification of mRNA from Y. lipolytica UOFS Y-0164
      grown on long-chain n-alkanes/fatty acids using ALK-1/ALK-REV
      to ALK-8/ALK-REV primer pairs
  3.2.8 Northern hybridizations
  3.2.9 Southern hybridizations
  3.2.10 Sequence analysis

Chapter Four: Results and Discussion

4.1 PCR amplification of CYP52 and Actin gene fragments from
    Y. lipolytica UOFS Y-0164

4.2 Cloning and sequencing of CYP52 gene fragments from Y. lipolytica
    UOFS Y-0164

4.3 Growth of Y. lipolytica UOFS Y-0164 on different hydrocarbons

4.4 Isolation of total RNA from Y. lipolytica UOFS Y-0164
4.5 RT-PCR amplification of *Y. lipolytica* UOFS Y-0164 *CYP52* gene fragments

4.6 The use of Northern hybridizations to monitor *CYP52* gene expression in *Y. lipolytica* UOFS Y-0164

4.7 The relationship between the evolutionary distance and regulation of the *Y. lipolytica* UOFS Y-0164 *CYP52* genes

Chapter Five: General Discussion

References

Summary

Samevatting
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACO</td>
<td>Fatty acyl CoA oxidase</td>
</tr>
<tr>
<td>AE</td>
<td>Sodium acetate-EDTA</td>
</tr>
<tr>
<td>AHR</td>
<td>Aryl hydrocarbon receptor</td>
</tr>
<tr>
<td>AHRE</td>
<td><em>cis</em>-acting aromatic hydrocarbon responsive element</td>
</tr>
<tr>
<td>AMV</td>
<td>Avian myeloblastosis virus</td>
</tr>
<tr>
<td>ARNT</td>
<td>Aryl hydrocarbon receptor nuclear translocator</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CAM</td>
<td>Camphor</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CYP</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>DBDs</td>
<td>DNA-binding domains</td>
</tr>
<tr>
<td>DEX</td>
<td>Dexamethasone</td>
</tr>
<tr>
<td>DME</td>
<td>Drug metabolizing enzyme</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DRE</td>
<td>Drug regulatory element</td>
</tr>
<tr>
<td>DRS</td>
<td>Distal regulatory element</td>
</tr>
<tr>
<td>EBP</td>
<td>Enhancer binding protein</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>FAD</td>
<td>Flavin adenine dinucleotide</td>
</tr>
<tr>
<td>GABP</td>
<td>y-Binding protein</td>
</tr>
<tr>
<td>GR</td>
<td>Glucocorticoid receptor</td>
</tr>
<tr>
<td>GRE</td>
<td>Glucocorticoid responsive element</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>HNF</td>
<td>Histone nuclear factor</td>
</tr>
<tr>
<td>HNF-4</td>
<td>Hepatocyte nuclear factor-4</td>
</tr>
<tr>
<td>hsp70</td>
<td>70kDa heat shock protein</td>
</tr>
<tr>
<td>hsp90</td>
<td>90kDa heat shock protein</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-1-thio-β-D-galactoside</td>
</tr>
<tr>
<td>KCl</td>
<td>Potassium chloride</td>
</tr>
<tr>
<td>LB</td>
<td>Lurie-Bertani medium</td>
</tr>
<tr>
<td>LBDs</td>
<td>Ligand binding domains</td>
</tr>
<tr>
<td>/cg</td>
<td>Luciferase chimeric gene</td>
</tr>
<tr>
<td>MC</td>
<td>3-Methylcholanthrene</td>
</tr>
<tr>
<td>MEF2</td>
<td>Multifunctional enzyme type 2</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium chloride</td>
</tr>
<tr>
<td>MgSO₄</td>
<td>Magnesium sulfate</td>
</tr>
<tr>
<td>MOPS</td>
<td>4-Morpholinepropanesulfonic acid</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger ribonucleic acid</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>Na₂PO₄</td>
<td>Sodium phosphate</td>
</tr>
<tr>
<td>NADH</td>
<td>Nicotinamide adenine dinucleotide</td>
</tr>
<tr>
<td>NADPH</td>
<td>Nicotinamide adenine dinucleotide phosphate</td>
</tr>
</tbody>
</table>
OP : Operator promoter
NF : Nuclear factor
NR : Nuclear receptor
NSAIDs : Non-steroidal anti-inflammatory drugs
ORE : Oleate response element
ORFs : Open reading frames
PAHs : Polycyclic aromatic hydrocarbons
PB : Phenobarbitone
PBREs : Phenobarbitone regulatory elements
PCI : Phenol/chloroform/iso-amyl alcohol
PCN : Pregnenolone 16α-carbonitrile
PCR : Polymerase chain reaction
perMEF2 : Peroxisomal multifunctional enzyme type 2
Pex : Peroxine
PKA : cAMP-dependent protein kinase
Pmp : Peroxisomal membrane protein
PP : Peroxisomal proliferator
PPAR : Peroxisomal proliferator activated receptor
PRC : Protein kinase C
PPRE : Peroxisomal proliferator responsive element
PRS : Proximal regulatory sequence
PTSs : Peroxisomal targeting signals
QC RT-PCR : Quantitative competitive reverse transcription-polymerase chain reaction
RAR : Retinoic acid receptor
RNA : Ribonucleic acid
RPA : Rnase protection assay
(RTDPN)PCR : Real time detection 5'-nuclease polymerase chain reaction
RT-PCR : Reverse transcription-polymerase chain reaction
RXR : Retinoid X receptor
SDI : Sex difference information
SDS : Sodium dodecyl sulfate
SSC : Sodium dodecyl sulfate-sodium chloride
ST : Sterigmatocystin
STAT : Signal transduction/activator of transcription
TAF : TATA box associating factor
TAO : Troleandomycin
Tag : Thermus aquaticus
TBP : TATA box binding protein
TCDD : 2,3,7,8-Tetrachlorodibenzo-p-dioxin
TE : Tris-EDTA
TfI : Thermus flavus
Tris : 2-Amino-2-(hydroxymethyl)-1,3-propanediol
UREs : Upstream regulating elements
X-gal : 5-Bromo-4-chloro-3-indolyl-β-D-galactoside
YEPO : Yeast extract/peptone/glucose
YNB : Yeast nitrogen base
List of Figures

Figure 2.1. Organization of the (+)-camphor oxidation pathway in P. putida, including the camDC4B operon and camR gene. OP, operator promoter (Aramaki et al., 1993).

Figure 2.2. Gene regulation of the Cytochromes P4501A by PAHs (Ramana & Kohli, 1998).

Figure 2.3. Gene regulation of the Cytochromes P4502B by Barbituates (Ramana & Kohli, 1998).

Figure 2.4. Trans-activation of the CYP2D5 promoter by C/EBPβ. (a) C/EBPβ does not bind or activate the CYP2D5 promoter. (b) The Sp1-binding site on the CYP2D5 promoter is functional. (c) With the presence of Sp1, C/EBPβ binds to the CYP2D5 promoter and increases the promoter activity. (Gonzalez & Lee, 1996).

Figure 2.5. Gene regulation of the Cytochromes P4503A by Glucocorticoids (Ramana & Kohli, 1998).

Figure 2.6. Activation of the CYP3A10 & CYP2C12 genes (Gonzalez & Lee, 1996).

Figure 2.7. Gene regulation of the Cytochromes P4504A by Peroxisome Proliferators (Ramana & Kohli, 1998).

Figure 4.1 A: Alignment of the 8 CYP52 genes present in Y. lipolytica. Eight specific primers (ALK-1 to ALK-8) were designed from the 5′-region of the CYP52 genes. The design of the common downstream primer (ALK-REV) was based on homologous nucleic acid sequences exhibited between all 8 CYP52 genes. Homology is indicated by asterisks and the primers by bold, italicized and underlined font. Alignment performed using CLUSTALW. B: Alignment of the Actin open reading frames (ORF) present in Rattus norvegicus, Homo sapiens, Gallus gallus, Candida apicola and Saccharomyces cerevisiae. Two conserved primers were designed: An upstream primer [ACT(F)] based on the 5′-region of the Actin ORFs as well as a downstream primer [ACT(R)]. Homology is indicated by asterisks and the primers by bold, italicized and underlined font. Alignment performed using CLUSTALW. C: PCR amplification of Y. lipolytica UOFS Y-0164 CYP52 and Actin gene fragments. Molecular size marker (lane 1): EcoRI/HindIII-digested λ phage ladder. Products of PCR reactions using primer pairs ALK-1/ALK-REV (lane 2); ALK-2/ALK-REV (lane 3); ALK-3/ALK-REV (lane 4); ALK-4/ALK-REV (lane 5); ALK-5/ALK-REV (lane 6); ALK-6/ALK-REV (lane 7); ALK-7/ALK-REV (lane 8); ALK-8/ALK-REV (lane 9); and ACT(F)/ACT(R) [lane 10].

Figure 4.2. Cloning of CYP52 Y. lipolytica UOFS Y-0164 PCR gene fragments. 1.5% (w/v) agarose gel containing 329bp EcoRI-digested pGEM®-T Easy Vector inserts. 1kb ladder (lanes 1 & 22); 329bp Y1ALK1 PCR gene fragments (lanes 2 & 3); Y2ALK2 (lanes 4 & 5); Y3ALK3 (lanes 6 & 7); Y4ALK4 (lanes 8 & 9); Y5ALK5 (lanes 10, 11 & 12); Y6ALK6 (lanes 13, 14 & 15); Y7ALK7 (lanes 16, 17 & 18) and Y8ALK8 (lanes 19, 20 & 21).

Figure 4.3. Utilization of n-alkanes/fatty acids by Y. lipolytica UOFS Y-0164. A: (◇) Control: (□) 2% pristane; (▲) 2% hexadecane; (●) 2% hexadecane & 2% pristane; (●) 2% octadecane; (◇) 2% hexadecane & 2% pristane. B: (□) 2% docosane; (▲) 2% docosane & 2% pristane; (●) 2% octacosane; (◇) 2% hexadecane & 2% pristane; (●) 0.5% Tween 80; (▲) 2% stearic acid & 0.5% Tween 80; (□) 2% behenic acid & 0.5% Tween 80.
Figure 4.4. RT-PCR amplification of 329bp *Y. lipolytica* UOFS Y-0164 CYP52 gene fragments. Cells were induced on (i) 2% pristane; (ii) 2% tetradeane; (iii) 2% hexadecane; (iv) 2% octadecane; (v) 2% docosane & 2% pristane; (vi) 2% octacosane & 2% pristane; (vii) 2% stearic acid & 0.5% Tween 80 and (viii) 2% behenic acid & 0.5% Tween 80. 100bp Molecular size marker (lane 1); Negative control (lane 2); ALK-1/ALK-REV primer pair (lane 3); ALK-2/ALK-REV primer pair (lane 4); ALK-3/ALK-REV primer pair (lane 5); ALK-4/ALK-REV primer pair (lane 6); ALK-5/ALK-REV primer pair (lane 7); ALK-6/ALK-REV primer pair (lane 8); ALK-7/ALK-REV primer pair (lane 9); ALK-8/ALK-REV primer pair (lane 10).

Figure 4.5. Southern hybridization analysis of *Y. lipolytica* UOFS Y-0164 CYP52 gene probes using CYP52 gene-specific PCR products. 329bp ALK-1/ALK-REV PCR product (lane 1); ALK-2/ALK-REV PCR product (lane 2); ALK-3/ALK-REV PCR product (lane 3); ALK-4/ALK-REV PCR product (lane 4); ALK-5/ALK-REV PCR product (lane 5); ALK-6/ALK-REV PCR product (lane 6); ALK-7/ALK-REV PCR product (lane 7); ALK-8/ALK-REV PCR product (lane 8).

Figure 4.6. Northern hybridization analysis of *Y/ALK1* gene expression after 5 hours of induction. Cells were induced on: 0.5% Tween 80 (lane 1); 2% pristane (lane 2); 2% tetradeane (lane 3); 2% hexadecane & 2% pristane (lane 4); 2% octadecane & 2% pristane (lane 5); 2% stearic acid & 0.5% Tween 80 (lane 6); 2% docosane & 2% pristane (lane 7); 2% octacosane & 2% pristane (lane 8).

Figure 4.7. Northern hybridization analysis of *Y. lipolytica* UOFS Y-0164 CYP52 gene expression after 5 hours. Cells were induced on: A: No C-source: Control (lane 1); 2% glucose (lane 2); 2% pristane (lane 3); 2% tetradeane (lane 4); 2% hexadecane (lane 5); 2% octadecane (lane 6); 2% docosane & 2% pristane (lane 7); 2% octacosane & 2% pristane (lane 8); 2% stearic acid & 0.5% Tween 80 (lane 9); 2% behenic acid & 0.5% Tween 80 (lane 10). B: 2% stearic acid & 0.5% Tween 80 (lane 1); 2% behenic acid & 0.5% Tween 80 (lane 2).

Figure 4.8. A: Growth of *Y. lipolytica* UOFS Y-0164 on 2% pristane (○); 2% octadecane (■); and 2% docosane & 2% pristane (▲). B: Northern hybridization analysis of *Y. lipolytica* UOFS Y-0164 CYP52 gene expression with time. Cells were induced on 2% pristane for 5 hours (lane 1), 9 hours (lane 2), and 13 hours (lane 3); 2% octadecane (lane 4), 9 (lane 5), and 13 hours (lane 6); and 2% docosane & 2% pristane for 5 (lane 7), 9 (lane 8), and 13 hours (lane 9), respectively.

Figure 4.9. Phylogenetic relationship exhibited between the *Y. lipolytica* UOFS Y-0164 CYP52 gene amino acid sequences. The phylogenetic tree was established on the basis of a multiple sequence alignment done with the CLUSTALW program (see Section 3.2.10). Despite the application of various combinations of reasonable input parameters, the phylogenetic tree did not differ significantly.
List of Tables

**TABLE 2.1.** A TABLE OF THE YEAST CYP52 GENES, THEIR GENE PRODUCTS AND THE SUBSTRATES THEY ARE INDUCED AND REPRESSED BY.

**TABLE 4.1.** COMPARISON OF THE NUCLEIC ACID *Y. lipolytica* CYP52 GENE AND CLONAL SEQUENCES
Chapter One

1.1 General Introduction

In the late 1940s and early 1950s the importance of oxidative reactions in the elimination of drugs in humans became apparent. The discovery that subcellular fractions from human liver catalyzed the oxidation of coumarin and biphenyl (Creaven & Williams, 1963) provided the enzymatic basis for this elimination process. The subsequent characterization of the drug-oxidizing system in the human liver (Kuntzman et al., 1966), showed activity to reside in the microsomal fraction and to require reduced nicotinamide adenine dinucleotide phosphate (NADPH). The concurrent year showed human liver fractions to be active in catalyzing many different routes of oxidation using chlorpromazine as a substrate (Coccia & Westerfield, 1967).

In 1964, the heme nature of a carbon monoxide-binding pigment (hemoprotein) located in the microsomal fraction of rodent liver was suggested to be of the b type (Omura & Sato, 1964). This porphyrin was then tentatively named cytochrome P450, meaning “a pigment with an absorption at 450nm in the reduced CO difference spectrum”. The nature of the heme prosthetic group has resulted in cytochromes being divided into four groups. The individual heme types, a, b, c and d have different side chains in their porphyrins, with well-established cytochromes being identified by subscript numerals with the letter indicating the groups, e.g., cytochrome a3.

Cytochrome P450 monooxygenases are the terminal oxidases of a number of biotransformation systems used by a variety of organisms. In addition to its oxygenase activity, terminal oxidases use molecular oxygen and electrons supplied by NADPH or NADH via a flavoprotein, to catalyze the NADPH-dependent reduction of molecular oxygen. The interaction of the hemoprotein with an electron donor, molecular oxygen, and a variety of organic substrates result in a wide diversity of oxidation products.

A cytochrome P450 monooxygenase often catalyzes the initial step in the degradation of lipophilic compounds such as n-alkanes and monoterpenes. Some cytochrome P450 genes are expressed constitutively while others, particularly those involved in xenobiotic metabolism, are inducible. In many cases, inducers are also substrates for the induced enzymes. P450 activities, for instance, remain elevated only as required. Enzyme induction usually enhances detoxification; thus, under most conditions, induction is a protective mechanism (Porter & Coon, 1991). Increased enzyme activity generally occurs due to an increase in the level of transcription (Gonzalez, 1989).

In recent years cytochrome P450 monooxygenases have become the subject of intensive research in many laboratories for three reasons. Firstly, from a fundamental point of view, the understanding of the versatility of this family of enzymes is of interest to those studying biological catalysis. Secondly, the elucidation of the reactions catalyzed by these enzymes have biomedical relevance for investigations into the fields of anesthesiology, endocrinology, nutrition, pathology, pharmacology, oncology and...
toxicology. Thirdly, the selective hydroxylation or side chain cleavage reactions of various organic compounds on a preparative scale, find application in synthetic organic chemistry. The functionalization of unactivated carbons is notoriously difficult to accomplish with chemical catalysts. The stereospecificity of such enzymatic reactions is put to use in the synthesis of chiral building blocks and steroid hydroxylations.

The efforts of a large number of individuals as well as the availability of new technologies have promoted the understanding of these complex enzymes. This being the case, areas demanding further attention include: (i) Three-dimensional analysis of additional protein structures encompassing their interactions with substrates. The need exists for crystallization of P450 membrane proteins due to the fact that they are too large for the application of global NMR methods. (ii) Additional analysis detailing the existing information concerning P450 catalytic mechanisms. (iii) Although much attention has been focussed on the transcription of individual P450 monoxygenases, the regulation of P450 enzymes is far more complex and requires more attention. (iv) There remain reactions of considerable importance that still require the isolation and characterization of the P450 monoxygenases involved. (v) Finally, the development of \textit{in vitro} as well as \textit{in vivo} systems detailing the influence individual P450 enzymes may have on various xenobiotics and endogenous chemicals that humans or other organisms might come into contact with.

1.2 Aims of the research project

The greatest utility of microbial systems can be seen in the arena of biodegradation work. Many bacterial and fungal P450s have already been characterized and are known to degrade certain chemicals of interest. Considerable potential exists for the isolation of new bacterial and fungal cytochrome P450 monoxygenases capable of oxidizing particular substrates. Generally, this could be done by identifying particular chemicals which can be utilized for growth by the P450-containing microorganisms.

In the early 1960s hydrocarbon biochemistry became a theme of industrial research. This led to research focussing on the biotechnological application of the alkane-assimilating yeasts like \textit{Candida maltosa}, \textit{Candida tropicalis}, and \textit{Yarrowia lipolytica}. The application of these yeasts in the biochemical synthesis of fatty acids, sterols, vitamins, and amino acids, as well as the production of single-cell protein as foodstuff using \textit{n}-alkanes as the carbon and energy source, had considerable commercial interest.

During recent years many genes coding for cytochrome P450 enzymes have been cloned from different organisms. In yeasts, the P450 genes that code for monoxygenases belong to multigene families, many of which have been isolated and characterized (Seghezzi et al., 1992; Zimmer et al., 1996; Lottermoser et al., 1996). Consequently, the differential expression of the individual P450 enzymes belonging to \textit{C. tropicalis}, \textit{C. maltosa} and \textit{C. apicola} paved the way for the identification of a multigene family in \textit{Y. lipolytica} (Iida et al., 2000). Of the eight P450 genes found to be present in \textit{Y. lipolytica} (Iida et al., 1998 & 2000), substrate specificities of the individual P450 forms are restricted to results obtained largely from gene disruptant experiments investigating growth on short-chain \textit{n}-
alkanes. This being the case, the involvement of the individual \textit{Y. lipolytica} P450 isoforms in long-chain \textit{n-}alkane/fatty acid assimilation had not been identified and is of considerable biotechnological interest.

The aims of this study therefore were:

i) To investigate the involvement of different cytochrome P450 monooxygenases in the hydroxylation of both short- and long-chain \textit{n-}alkanes and fatty acids by \textit{Y. lipolytica}.

ii) To monitor the differential expression of the eight \textit{Y. lipolytica} P450 genes.

iii) To gain more insight into the P450s structure-function relationship in an evolutionary context.

\textbf{Chapter Two: The regulation of Cytochrome P450 monooxygenases in prokaryotic and eukaryotic organisms – a literature overview}

\textbf{2.1 Introduction.}

The mitochondrial and microsomal multisubstrate mixed function cytochrome P450 monooxygenase system has been intensively studied for the past 30 years. The evolution of this system can be attributed to the constant exposure of organisms to xenobiotics. Early studies recognized that many chemicals were able to induce their own metabolism, and this added to the complexity of NAD(P)H/O\textsubscript{2}-dependent oxidation. An understanding of the molecular biology of P450s has been greatly aided by the tremendous progress in the ‘70s and early ‘80s in purifying a number of P450 forms. Such protein purification efforts identified the existence of multiple P450 forms present in mammals, other eukaryotes as well as some prokaryotes.

The genes that constitute the cytochrome P450 superfamily code for a large variety of enzymes that exhibit the following characteristics: i) they contain a noncovalently bound haem group; ii) substrates are oxygenated by an oxygen atom derived from atmospheric oxygen, using reducing equivalents from NADPH or NADH; iii) a second enzyme functions to transfer the reducing equivalents to the P450; and iv) with a few exceptions in bacteria, they are intrinsic membrane proteins bound firmly to intracellular membranes.

cDNA and gene cloning in plants, fungi, bacteria and mammals have demonstrated the presence of hundreds of P450s. Among the xenobiotic-metabolizing P450s are found the \textit{CYP1}, \textit{CYP2}, \textit{CYP3} and \textit{CYP4} families. Marked differences have been noted in the expression of different P450s present in separate subfamilies. The prevalence of genetic polymorphisms involved in the expression of different P450s is also a common occurrence. Molecular mechanisms of gene regulation are used to study and fully understand the differences in P450 expression found between various species. In higher organisms transcriptional regulation contributes largely to the control of tissue and
developmental stage-dependent gene expression. The accessibility of the chromatin structure at the gene locus and the interaction of transcription factors with their corresponding cis-acting regulatory elements, govern the transcriptional activity of a gene. Ligand-independent expression factors govern the constitutive expression of genes. Some of these factors are ubiquitously expressed in many different tissues or cell types, while others appear to be more restrictively expressed. Up-regulation of certain constitutively transcribed genes, brought about by ligand-dependent transcription factors, have been found to belong to the nuclear receptor superfamily (Mangelsdorf et al., 1995), or the basic helix-loop-helix/PAS domain superfamily (Swanson & Bradfield, 1993). The in situ formation of metabolites due to hormonally regulated extra-hepatic P450s in specific tissues can influence target organ toxicity, sensitivity and responsiveness. This provides the basis for the characterization of P450 forms in various tissues as well as the fluctuations in their levels under physiological conditions. Such studies can elucidate mechanisms of hormonal P450 gene expression as well as shedding light on the individual steps in the signal transduction pathways. The induction of subsets of P450s by individual compounds such as polycyclic hydrocarbons, glucocorticoids, barbiturates, peroxisome proliferators, and ethanol is as a result of an increase in gene transcription with concomitant increases in the amounts of P450s present in the cell (Gonzalez, 1989).

To date, the molecular biology approach has offered a valuable alternative to a biochemical or immunological P450 research approach. It is now possible to isolate, identify, and characterize a large variety of P450s in a range of tissues and organisms using molecular cloning techniques. The use of molecular cloning is seen to play a large part in studying the catalytic specificities of various P450s, structure-function relationships found within the catalytic cycles of P450s, as well as the isolation of different P450 genes used to investigate P450 gene regulation.

The mammalian xenobiotic metabolizing P450 monooxygenases have been classified in the first four gene families, CYP1-4. They exhibit broad, but overlapping, substrate and product specificities and their expression is species-, sex- and tissue-specific (Nebert et al., 1993; Gonzalez, 1989; Kemper, 1998; Lund et al., 1991). Several liver specific transcription factors involved in the regulation of P450 expression at post-transcriptional and/or post-translational levels have been identified (Honkakoski & Negishi, 2000). The study of bacterial P450s with respect to the regulation of their expression was made possible due to mechanistic relationships analogous to mammalian cytochrome P450s (Fulco, 1991; He & Fulco, 1991; Wen & Fulco, 1987; Narhi & Fulco, 1982). The regulation of fungal P450s, however, remains to be investigated for the major part of the fungal P450 genes thus far identified (Nelson et al., 1993). The fungal P450 monooxygenases, classified in the gene families CYP51-66, code for enzymes induced by and involved in the assimilation of a range of substrates varying from sterols to hydrocarbons. The observation that many yeasts are able to grow on n-alkanes as sole carbon and energy source, resulted in the extensive study of the hydrocarbon assimilating yeasts of the genus Candida. In these yeasts cytochrome P450 was induced by the hydrocarbon substrate and the corresponding genes coding for the enzymes were classified within the CYP52 gene family (Nelson et al., 1993). The isolation and characterization of P450s involved in hydrocarbon assimilation in the Candida spp.,
resulted in the identification of the CYP52 multigene families found in yeasts such as C. tropicalis, C. maltosa, C. apicola and Y. lipolytica (Seghezzi et al., 1992; Zimmer et al., 1996; Lottermoser et al., 1996; Iida et al., 1998 & 2000). Additional CYP genes isolated from Saccharomyces cerevisiae, Nectria haematococca, Aspergillus nidulans, Rhodotorula minuta, Fusarium oxysporum, Fusarium sporotrichioides and Phanerochaete chrysosporium, among others, are not directly involved in hydrocarbon assimilation (Kullman & Matsumura, 1997; Hohn et al., 1995; Kizawa et al., 1991; Fujii et al., 1989; Kelkar et al., 1997; Maloney & VanEtten, 1994) and have been classified in other fungal P450 gene families.

2.2 The P450 gene superfamily

2.2.1 Summary of the P450 gene nomenclature

The isolation of the first individual P450 enzymes, named according to their catalytic activity towards a particular substrate, resulted in the same enzyme being given several different names in different laboratories. As a result, a P450 classification system was recommended that grouped these P450s into a gene superfamily, that was further subdivided into gene families and later into gene subfamilies (Nebert et al., 1991). Cytochrome P450 enzymes in all species were designated using the prefix CYP, excluding that of Cyp for the mouse. Members constituting a particular gene family had to have at least 40% homology between the corresponding amino acid sequences. These gene families are designated using an Arabic number (e.g. CYP2). A subfamily of the gene is denoted by a capital letter, followed by an Arabic numeral corresponding to the individual gene, e.g. Cyp2D1 (Slaughter & Edwards, 1995; Nebert et al., 1991).

The long-established overlapping substrate specificities of P450s represent a feature that is in keeping with the similar catalytic activities exhibited between P450s present in separate subfamilies. As a result, the nomenclature system developed by Nebert et al. is not dependent on P450 catalytic activities or function. The assignment of orthologous genes found within species is another inherent problem in P450 nomenclature. Many P450 genes within a particular species have been scrambled due to gene conversion events. Such gene conversion events within individual species have given rise to unique P450 genes, thereby complicating the nomenclature of the P450 gene superfamily.

The emergence of the cytochrome P450 gene superfamily is as a result of gene duplications followed by divergences due to mutations from an ancestral gene that existed more than two billion years ago (Nebert & Gonzalez, 1987). Orthologous gene assignments between species has become impossible due to the numerous gene conversion and duplication events found in several subfamilies. As a result, the degree of inter- and intra-species gene variations observed within the P450 gene superfamily, comprising 74 gene families with more than 500 genes, is very high. For widely diverged species and subfamilies containing three or more genes, the classification of proteins encoded by orthologous genes are usually numbered sequentially on a chronological basis.
2.2.2. Regulation of the P450 gene superfamily

Families or subfamilies of genes were initially thought to be regulated in much the same way as 3-methylcholanthrene-inducible genes or phenobarbital clusters. However, it was later realized that a combination of transcriptional activation, post-transcriptional, and post-translational regulation events were responsible for P450 “induction” (Gonzalez, 1989; Nebert & Gonzalez, 1987). Such P450 up-regulation by a specific class of inducers need not be restricted to any specific gene family or subfamily. Polycyclic hydrocarbons and phenobarbital are known to induce mammalian enzymes coded for by the CYP1A1, CYP1A2, as well as the CYP2B, CYP2C and CYP3A genes, respectively. Phenobarbital has also been reported to be involved in the induction of the bacterial CYP102 and CYP106 genes, as well as the mouse Cyp2b10 and chicken CYPH1 genes (Kemper, 1998). Hormonal and metalloregulation of gene expression is also seen for CYP2A4, CYP2A5 as well as CUP1, CRS5 and SOD1 located in the mouse and Saccharomyces cerevisiae, respectively (Lund et al., 1991; Winge, 1998).

The regulation of P450 gene expression at either post-transcriptional and/or post-translational levels upon exposure to xenobiotics has been observed for the rabbit (Muerhoff et al., 1992a), the rat (Hardwick et al., 1983), and the mouse (Honkakoski et al., 1996). The induction of different P450 isoforms by a variety of xenobiotics presents a unique model to investigate the regulation of gene expression, as these genes are found in both prokaryotic and eukaryotic organisms alike.

2.3. Bacterial Cytochrome P450 monoxygenases

Prokaryotic cytochrome P450 systems utilize two protein components, viz, an iron sulphur protein and an FAD-containing reductase. The iron sulphur protein shuttles between the reductase and the cytochrome P450, transferring electrons to the cytochrome P450 one at a time. Prokaryotic P450 systems render themselves more amenable to study due to the fact that eukaryotic P450 systems are difficult to obtain in large quantities of highly pure, homogeneous forms for structure-function studies. Complicating factors that restrict the study of eukaryotic cytochrome P450 enzymes are the fact that they are membrane-associated, and the various isoforms have very similar physical properties.

To date, bacterial cytochrome P450 monoxygenases have been found to be involved in the hydroxylation of monoterpenes, saturated and unsaturated fatty acids, amides, alcohols, and barbituates. These enzymes include the (+)-camphor monoxygenase (P450cam/CYP101) from Pseudomonas putida (Rheinwald et al., 1973; Unger et al., 1986), the fatty acid/barbiturate monoxygenase (P450BAF/CYP102) from Bacillus megaterium (Narhi & Fulco, 1987; Wen & Fulco, 1987), the α-terpineol monoxygenase (CYP108) from a Pseudomonas sp. (Peterson et al., 1992), and a linalool monoxygenase (P450lin/CYP111) from P. incognita (Ropp et al., 1993). Other cytochrome P450 monoxygenases isolated from bacteria, and involved in the oxidative metabolism of both endogenous and exogenous hydrophobic compounds include: cytochrome P450pinF1 isolated from Agrobacterium tumefaciens (the product of the CYP104 gene); cytochrome
P450SU1 and cytochrome P450SU2 isolated from *Streptomyces griseolus* ATCC 11796 (the products of the CYP105A1 and CYP105B1 genes, respectively); cytochrome P450chop isolated from another *Streptomyces* sp. (the product of the CYP105C1 gene); and cytochrome P450sub isolated from *Bacillus subtilis* (the product of the CYP109 gene) [Nebert et al., 1991].

The involvement of naturally occurring plasmids in the degradation of many aromatic, haloaromatic, aliphatic, and various acyclic isoprenoid compounds has been described (Rheinwald et al., 1973; Vandenbergh & Wright, 1983; Peterson et al., 1992; Van Beilen et al., 1994; Colocousi et al., 1996). Such alkane/monoterpene-oxidizing plasmid-encoded genes have been found in *P. oleovorans* and *P. putida*, respectively (Van Beilen et al., 1994; Rheinwald et al., 1973; Vandenbergh & Wright, 1983). Chromosomal-encoded genes governing the hydroxylation of various monoterpines have also been reported for *P. incognita*, *P. florescens*, and other *P.* spp. (Ropp et al., 1993; Colocousi et al., 1996; Peterson et al., 1992).

### 2.3.1 The *Pseudomonas putida* Cytochrome P450cam

The camphor 5-exo-monoxygenase from *P. putida* is an inducible, multicomponent P450 system that incorporates two additional redox proteins, an FAD reductase, and a Fe₂S₂Cys₄ redoxin, transferring electrons from NADH to P450cam. In a complete reconstituted system, P450cam catalyzes the stereospecific 5-exo hydroxylation of a bicyclic terpene, camphor, utilized as a sole carbon source. Crystallography was used to determine the structure of (+)-camphor monoxygenase to a resolution of 2.6Å (Poulos et al., 1985). The (+)-camphor crystal structure was subsequently used together with mutagenic, biochemical and biophysical studies to allow an in depth analysis into the problem of how P450s activate molecular oxygen, control stereoselectivity, and transfer electrons (Poulos et al., 1985; Liu et al., 1995; Kadkhodayan et al., 1995; Stevenson et al., 1996).

Cytochrome P450cam, the camC gene product from the *P. putida* PpG1 CAM plasmid, has provided the model for physical and chemical studies of hydroxylase systems. The genes initially involved in camphor oxidation constitute a monoxygenase system comprising a 45kDa NADH-putidaredoxin reductase (encoded by the camA gene), a 12kDa putidaredoxin (encoded by the camB gene), and a 47kDa cytochrome P450cam (encoded by the camC gene) [Koga et al., 1986]. The resulting alcohol is converted to 2,5-diketocamphane by the 80kDa 5-exo-hydroxycamphor dehydrogenase (encoded by the camD gene) [Fig. 2.1]. Collectively, all four genes form an operon called the cytochrome P450cam hydroxylase operon (camDCAB) [Koga et al., 1986 & 1989].
The camDCAB operon is under negative control by camR, located immediately upstream of the camD gene. camR is transcribed in the opposite direction to that of the camDCAB operon (Fujita et al., 1993; Koga et al., 1986), and shows maximal expression in the presence of (+)-camphor (Koga et al., 1986).

Aramaki et al. (1993) proposed a molecular mechanism governing the regulation of the camDCAB and camR genes in P. putida. Both the camR and camDCAB genes appear to have a common functional operator that represses the transcription of both genes. In the absence of camphor repression is mediated by the CamR protein. Induction of both genes is brought about by the addition of camphor, releasing the repressing protein CamR. In vitro transcription assays showed that camR is transcribed by the σ70 RNA polymerase, while camDCAB is not (Fujita et al., 1993). It was suggested that transcription from the camDCAB promoter could be mediated by an alternative polymerase, or, by various positive factor(s) that might interact with the σ70 RNA polymerase (Fujita et al., 1993).

2.3.2 The Bacillus megaterium Cytochrome P450_{BM,3}

Cytochrome P450_{BM,3} from B. megaterium is a soluble P450-dependent monoxygenase that catalyzes the (ω-2)-hydroxylation of saturated long-chain fatty acids, amides, alcohols (Ho & Fulco, 1976; Matson et al., 1977), and the hydroxylation and/or epoxidation of unsaturated fatty acids (Ruettinger & Fulco, 1981). Cytochrome P450_{BM,3} is also inducible by peroxisome proliferators, a characteristic of the fatty acid monoxygenases belonging to the CYP44 gene family (Gonzalez, 1989), as well as by the non-steroidal anti-inflammatory drugs (NSAIDs) ibuprofen, ketoprofen, and indomethacin (English et al., 1996). This NSAID-induced P450_{BM,3} incorporates both a
P450 and an NADPH:P450 reductase, constituting two proteolytically separable functional domains, thereby comprising a single 119kDa polypeptide (Ruettinger et al., 1989; Narhi & Fulco, 1987). Both the P450 and reductase domains define new gene families. Here, the P450 exhibits a 25% sequence similarity with the fatty acid ω-hydroxylases belonging to the CYP4A gene family, while the reductase shares a 33% sequence similarity with the mammalian liver NADPH:P450 reductases (Ruettinger et al., 1989).

The cloning of the P450BM-3 gene into *Escherichia coli* by Wen & Fulco. (1987), demonstrated a high level of expression of functional P450BM-3 protein using an immunochemical screening technique. Expression of the P450BM-3 gene was found to be directed by a promoter included in the DNA insert (Wen & Fulco, 1987). However, insertion of the cloned gene on a shuttle vector into *B. megaterium* resulted in a low basal expression with a dramatic pentobarbital-induction response, while synthesis proceeded constitutively at a high rate in *E.coli*, and was not induced by pentobarbital. It was initially thought that different promoters were responsible for the expression of the P450BM-3 gene in *E. coli* and *B. megaterium*, but work done by Ruettinger et al. (1989) showed that the transcription initiation sites of the P450BM-3 gene to be identical in both organisms, and that equivalent transcripts were produced. The P450BM-3 gene appears to be differentially expressed in both *E. coli* and *B. megaterium*. This could be attributed to differences in either the σ-subunits of the RNA polymerases of the two organisms, or the presence of specific trans-acting elements (repressor and/or enhancer proteins) present in *B. megaterium* and not in *E. coli*. Transcriptional activation of the P450BM-3 gene in *B. megaterium* is under positive control and is mediated by the binding of a protein to at least two regulatory domains, namely R1 and R2 (Wen et al., 1989). The role of barbituates in P450BM-3 gene induction is hypothesized to either facilitate the binding of this regulatory protein or to enhance its level in the cell by mediating either an increase in its rate of synthesis or by retarding its breakdown.

2.3.3 Other bacterial Cytochrome P450 monooxygenases

Cytochrome P450s involved in the utilization of linalool and α-terpineol as a sole carbon source by *Pseudomonas incognita* and another *Pseudomonas* sp., respectively, has been previously described (Ropp et al., 1993; Peterson et al., 1992). The *Pseudomonas* sp. P450, designated P450lin/CYP111, is best characterized in terms of biophysical phenomena. The *Pseudomonas incognita* P450, P450Lin/CYP101, has not been as extensively characterized (Nebert et al., 1991). The biophysical characterization of P450lin/CYP111 includes electrophoretic mobilities, cursory spectral characterizations and heterologous reconstitution assays. Although the amino acid sequence alignments of P450lin/CYP111 and P450Lin/CYP101 only showed a 25% similarity, the conservation of amino acids involved in O2, heme binding and residues involved in hydrogen bond interactions with heme propionates, was observed (Ropp et al., 1993). In 1992 a novel cytochrome P450 system, designated cytochrome P450terp, was isolated and purified from a new *Pseudomonas* sp. (Peterson et al., 1992). Sequence analysis of clones believed to
contain the gene encoding cytochrome P450\textsubscript{terp} revealed the presence of five open reading frames (ORFs). The ORFs were found to include i) an alcohol dehydrogenase ii) an aldehyde dehydrogenase iii) a cytochrome P450 iv) a terp reductase, and v) a terp (Peterson \textit{et al.}, 1992). It was concluded that the complete \textit{terp} operon had been identified. Sequence comparisons with other cytochromes P450 revealed it to be the first member of the \textit{CYP108} gene family (Peterson \textit{et al.}, 1992). The preliminary crystallization and x-ray diffraction analysis of P450\textsubscript{terp} and the hemoprotein domain of P450\textsubscript{B3.1-3} (Bodupalli \textit{et al.}, 1992), paved the way for the refinement of P450\textsubscript{terp} at 2.3\AA resolution (Hasemann \textit{et al.}, 1994). Further studies detailed the relationship of active site topology to substrate specificity for cytochrome P450\textsubscript{terp} (Fruetel \textit{et al.}, 1994).

\section*{2.4. Mammalian Cytochrome P450 monooxygenases}

Mammalian P450s can be divided into two major classes based on the enzyme from which they receive electrons, and on their intracellular location. The first class constitutes enzymes involved in steroid synthesis, \textit{viz} steroid 11\beta-hydroxylase-side-chain cleavage enzymes and mitochondrial P450s, and are located in the mitochondria of the adrenal cortex. Membrane-free polyribosomes synthesize these enzymes (Nabi \textit{et al.}, 1983) as large precursors (DuBois \textit{et al.}, 1981) which are then transported into the organelle concomitant with the cleavage and removal of an NH\textsubscript{2}-terminal extrapeptide (Morohashi \textit{et al.}, 1984; DuBois \textit{et al.}, 1981). The iron sulphur protein adrenodoxin transfers electrons to the mitochondrial P450s via NADPH-adrenodoxin reductase. The majority of P450 monooxygenases are bound up in the endoplasmic reticulum membrane and constitute the second class of mammalian P450 enzymes. Membrane-bound polyribosomes synthesize these enzymes (Gonzalez \& Kasper, 1980), which are then directly inserted into the lipid bilayer via a signal recognition system (Bar-Nun \textit{et al.}, 1980). A flavoprotein NADPH-P450 oxidoreductase transfers the electrons to the endoplasmic reticulum-bound P450 enzymes.

The P450 monooxygenases found within the \textit{CYP1}, \textit{CYP2} and \textit{CYP3} gene families, the major mammalian P450 gene families, metabolize carcinogens and drugs. These enzymes are also known to hydroxylate steroids such as progesterone, estrogen and testosterone. The \textit{CYP4} mammalian gene family codes for fatty acid hydroxylases responsible for metabolizing lauric acid, arachidonic acid, palmitic acid and prostaglandins.

\subsection*{2.4.1 Aromatic-hydrocarbon-inducible Cytochrome P450 monooxygenases}

\subsection*{2.4.1.1 The P450I gene family}

The P450I (\textit{CYP1}) gene family was previously thought to be composed of two members, designated \textit{CYP1A1} and \textit{CYP1A2} (Gonzalez, 1989). Both \textit{CYP1A1} and \textit{CYP1A2} are detected following treatment with inducers such as 2,3,7,8-tetrachlorodibeno-p-dioxin (TCDD) and/or 3-methylcholanthrene (MC), even though \textit{CYP1A2} is constitutively expressed in the liver. \textit{CYP1A1} and \textit{CYP1A2} have high catalytic activities in the presence of benzo[a]pyrene, an aromatic hydrocarbon, and acrylamines, respectively (Sakaki \textit{et al.}, 1984; Guengerich \textit{et al.}, 1982; Kuwahara \textit{et al.}, 1984; Goldstein \textit{et al.}, 1982; Johnson &
Muller-Eberhard, 1977; Johnson et al., 1980). CYPIA1 and CYPIA2 are ubiquitous in mammals and have similar catalytic activities towards a variety of xenobiotics. A new member of this family, a novel human cytochrome P450 designated P450B1 (CYPIBI), represents a second subfamily (Sutter et al., 1991). Although CYPIA1, CYPIA2 and CYPIBI most likely evolved from the same ancestral gene (Sutter et al., 1994), CYPIA1 and CYPIA2 map to human chromosome 15 (Hildebrand et al., 1985) while CYPIBI maps to human chromosome 2 (Sutter et al., 1994; Tang et al., 1996). CYPIBI is constitutively expressed in the testes, adrenals and ovaries and is inducible by adrenocorticotropin, peptide hormones, planar aromatic hydrocarbons and TCDD (Otto et al., 1991 & 1992; Tang et al., 1996; Sutter et al., 1994).

2.4.1.2 Regulation of the CYPI gene family

Polycyclic aromatic hydrocarbons (PAHs), such as the carcinogens MC, TCDD and benzo[a]pyrene, are typical inducers of several P450s, most notably CYPIA1, CYPIA2 and CYPIBI (Guengerich, 1991; Coon et al., 1992; Gonzalez & Lee, 1996; Tang et al., 1996). Studies performed using cycloheximide and actinomycin D, inhibitors of RNA and therefore protein synthesis, suggested that CYPIA1 induction and expression was regulated at both the transcriptional and post-transcriptional levels. The induction of cytochrome P450s by PAHs, specifically that of CYPIA1, has been extensively studied. Poland et al. (1976 & 1982) showed that the mouse liver contained a protein that bound TCDD saturably, reversibly, and with a high affinity indicating the functional ligand-binding properties of a receptor. Later this protein was named the aryl hydrocarbon receptor (AHR) due to the fact that it also binds MC and benzo[a]pyrene.

Regulation of the CYPI gene family is proposed to be mediated via the AHR protein (Poland et al., 1976 & 1982; Perdew et al., 1988; Ramana & Kohli, 1998). In the uninduced state AHR, a member of the helix-loop-helix/PAS family of transcription factors, forms a binary complex with the 90kDa heat shock protein (hsp90) [Fig. 2.2]. The configuration of the AHR-hsp90 binary complex dissociates upon the binding of a PAH, with concomitant release of hsp90 (Perdew, 1988). The binding of PAH to AHR results in the formation of a new binary complex PAH-AHR, which then translocates to the nucleus where it binds AHR nuclear translocator (ARNT). This ternary complex, PAH-AHR-ARNT, then binds cis-acting aromatic hydrocarbon responsive element (AHRE) thereby increasing the transcription of the CYPIA1 gene (Ramana & Kohli, 1998)
The presence of at least three cis-acting regulatory regions responsible for inductive expression relative to the transcriptional start site were identified by Sogawa et al. (1986) using external deletion analysis. The regions identified were from nucleotide -3674 to -3067, -1682 to -1429, and -1139 to -1029. The region most important in CYP1A inducibility, termed the drug regulatory element (DRE), is the latter 10bp regulatory element. The DRE and its homologues are tandemly arranged within this region, with a consensus sequence \( 5'\text{-G}^\text{C}\text{N}^\text{N}^\text{C}\text{GCTGG-3'} \) (Sogawa et al., 1986).

Studies performed in the past have identified the role of signal transduction by TCDD in the regulation of target genes, specifically CYP1A gene expression by PAHs. It has been suggested that the induction of CYP1A gene expression by TCDD may require a protein kinase C (PRC)-dependent pathway (Berghard et al., 1993). It was shown, \textit{in vitro}, that the DNA-binding activity of the ligand-activated TCDD receptor was inhibited by dephosphorylation. The TCDD receptor function appears to be regulated by a complex pattern of phosphorylations, as ARNT appeared to require phosphorylation to interact with the receptor. The binding of the PAH-AHR-ARNT ternary complex within the major groove of the double helix (Carrier et al., 1992), as well as the Ah receptor-induced DNA distortion of the AHRE (Elferink & Whitlock, 1990), are factors that have been proposed to mediate the loss of the nucleosomal structure at the promoter. Once this has occurred, the constitutive transcription factors, needed for transcription of the CYP1A1 gene, can bind and enhance transcription.

Where an increase in the transcription mainly regulated CYP1A1 gene expression, post-transcriptional mRNA stabilization appears to increase CYP1A2 gene expression. However, the mechanisms of CYP1A2 gene regulation are not completely understood. The distal regulatory sequences (DRS) of the CYP1A2 gene has been recently found to
contain cis-acting positive and negative elements at -2218 to -2187 and -2124 to -2098 nt, respectively (Chung & Bresnick, 1997; Ramana & Kohli, 1998).

Work done recently by Tang et al. (1996) on the mapping of the human CYP1B1 gene pinpointed the location of the gene at 2p21-22 on the human chromosome 2, a region known to be involved in development, signal transduction, and hormone responsiveness. Nucleotide sequence analysis of the 5'-flanking region of the CYP1B1 gene identified the presence of 4DREs in the first intron (of two), three in the second exon (of three), and five in the second intron. Sutter et al. (1991) showed that TCDD caused a 2-3-fold increase in the rate of transcription of CYP1B1. Tang et al. (1996) were able to identify a TCDD-responsive region containing three DREs, found between -1022 to -835 nt, in the CYP1B1 gene. They propose that at least one of the three DREs may interact weakly with activated AHR-ARNT complex, thereby promoting the enhancement of CYP1B1 expression through a synergic effect with certain transcription factors (Tang et al., 1996).

2.4.2 Phenobarbital-inducible Cytochrome P450 monooxygenases

2.4.2.1 The P450II gene family

The mammalian P450II (CYP2) gene family consists of seven subfamilies, constitutively expressed in the liver. The CYP2 gene subfamilies are transcriptionally activated at distinct stages of development, with specific P450 genes being preferentially expressed in a particular sex. Different liver-enriched transcription factors, including histone nuclear factor-1α (HNF-1α), HNF-3, HNF-4, C/EBPβ, as well as the more ubiquitously expressed factors Sp1, GABPα/β and NF2d9 give rise to the diversity exhibited in the transcriptional mechanisms of liver-specific CYP2 gene expression (Gonzalez & Lee, 1996).

2.4.2.1.1 The CYP2A gene subfamily

The CYP2A gene subfamily has been characterized at the protein level following extensive studies performed on the rat. The first P450 of this subfamily to be purified, designated CYP2A1, specifically hydroxylates testosterone at the 7α position (Nagata et al., 1987; Waxman et al., 1983). Hydroxylation of testosterone at the 6α position occurs to a lesser extent, and is therefore a minor metabolite. A more relaxed specificity toward testosterone hydroxylation is seen when examining the second gene of the CYP2 gene subfamily, namely CYP2A2 (Matsunaga et al., 1988; Jansson et al., 1985). Hydroxylation of testosterone by CYP2A2 produces 2α, 6β, 7α, 15α, 15β and 16α hydroxylated metabolites.

The cDNA for CYP2A1, CYP2A2 and a third gene in the subfamily, designated CYP2A3, was isolated from the rat and sequenced (Nagata et al., 1987; Matsunaga et al., 1988; Kimura et al., 1989). An 88% cDNA-deduced amino acid sequence similarity exhibited between CYP2A1 and CYP2A2 suggests the occurrence of a gene conversion event in the rat CYP2A gene subfamily (Matsunaga et al., 1988). The third CYP2A gene subfamily member, CYP2A3, showed amino acid sequence similarities of 71% and 73% when
compared to \textit{CYP2A1} and \textit{CYP2A2}, respectively (Kimura et al., 1989). It has been suggested that \textit{CYP2A3} diverged from a common \textit{CYP2A1/CYP2A2} ancestor.

An active testosterone 15α-hydroxylase, designated P450_{15α}, was cloned and sequenced from the mouse. Two related cDNAs, termed 15α type I and type II, were found to exhibit a 98% cDNA-deduced amino acid sequence similarity. When comparing the latter cDNA-deduced amino acid sequences to those of the \textit{CYP2A1}, \textit{CYP2A2} and \textit{CYP2A3} genes, similarities of 70%, 75%, and 90% were obtained, respectively. It was suggested that the 15α type I and type II counterparts are orthologous to the rat \textit{CYP2A3} gene (Kimura et al., 1989).

A partial human P450 cDNA was isolated, sequenced and designated P450(I) as part of the \textit{CYP2A} gene subfamily (Phillips et al., 1985). Miles et al. (1989) and Yamano et al. (1989a) later isolated and sequenced a cDNA coding for the entire P450(I) human gene. It was found that the human enzyme was orthologous to the \textit{CYP2A3} gene as it displayed an 85% deduced amino acid sequence with the rat \textit{CYP2A3} gene, while only having a 69% and 65% deduced amino acid sequence similarity with the \textit{CYP2A1} and \textit{CYP2A2} genes, respectively (Miles et al., 1989; Yamano et al., 1989b).

\subsection{2.4.2.1.2 The \textit{CYP2B} gene subfamily}

Two \textit{CYP2B} genes, designated \textit{CYP2B1} and \textit{CYP2B2}, have been extensively studied in the rat and are noteworthy due to the fact that they are induced by phenobarbital (Guengerich et al., 1982; Waxman & Walsh, 1982). The \textit{CYP2B1} enzyme has a 2- to 10-fold higher level of activity, depending on the substrate, than the \textit{CYP2B2} enzyme, although they exhibit similar broad yet overlapping substrate specificities. Here, \textit{CYP2B1} and \textit{CYP2B2} exhibit a 97% cDNA-deduced amino acid sequence similarity (Fujii-Kuriyama et al., 1982; Kumar et al., 1983; Suwa et al., 1985). A third P450 belonging to the \textit{CYP2B} gene subfamily was identified through cDNA cloning, and designated \textit{CYP2B3} (Gonzalez, 1989). Upon comparison of the \textit{CYP2B3} cDNA-deduced amino acid sequence with those corresponding to the \textit{CYP2B1} and \textit{CYP2B2} genes, percentages point towards an approximately 75 million year old ancestral gene. Two \textit{CYP2B} genes have also been identified in the rabbit and exhibit an approximate 95% amino acid sequence similarity (Komori et al., 1988).

\subsection{2.4.2.1.3 The \textit{CYP2C} gene subfamily}

The genes belonging to the \textit{CYP2C} gene subfamily are noted for their developmental and sex-specific expression in rats, and are generally thought to represent a class of constitutively expressed genes. The \textit{CYP2C} gene subfamily comprises two enzymes expressed in both males and females (P450PB-1 and P450f), an adult female-specific P450 (P450f/15β) and two adult male-specific P450s (P450h/16α and P450g). The \textit{CYP2C} enzymes have been found to exhibit broad overlapping specificities towards a variety of chemicals, while some show high levels of catalytic activity towards steroids (Gonzalez, 1989).
A number of rat cDNAs have been characterized. These include cDNAs coding for P450PB-1 (CYP2C6), P450f (CYP2C7) [Gonzalez et al., 1986; Friedberg et al., 1986; Kimura et al., 1988], the female-specific P450i/15β (CYP2C12) [Zaphiropoulos et al., 1988], and the male-specific P450h/16α (CYP2C11) [Yoshioka et al., 1987; Zaphiropoulos et al., 1988]. At least seven P450s appear to be expressed in the rabbit. Those characterized comprise progesterone 21-hydroxylase (Tukey et al., 1985), four phenobarbital-induced forms from a liver library designated CYP2C to CYP2C4 (Leighton et al., 1984), the P450 1 (CYP2C5), as well as pHP3 (CYP2C14) and b32-3 (CYP2C15) [Imai et al., 1988].

CYP2C8 and CYP2C9 are two P450 cDNAs isolated and sequenced from human liver libraries (Okino et al., 1987; Kimura et al., 1987; Umbenhauer et al., 1987). The CYP2C9 cDNA sequence corresponds to two enzymes isolated from the human liver, namely P450_{HPI} and P450_{MP2} (Shimada et al., 1986).

2.4.2.1.4 The CYP2D gene subfamily

Rat and human P450 enzymes belonging to the CYP2D subfamily were purified (Larrey et al., 1984; Distlerath et al., 1985; Gut et al., 1986), and were found to be involved in the oxidation of the drugs bufuralol and debrisoquine. Immunological and cDNA cloning studies performed on the rat identified the presence of five CYP2D genes (Gonzalez et al., 1989; Matsunaga et al., 1989). The five genes were found to have between 75% and 95% deduced-amino acid sequence similarity, and were designated CYP2D1 to CYP2D5, respectively (Gonzalez et al., 1989; Matsunaga et al., 1989). A polymorphism found in humans associated with the inability to be able to hydroxylate debrisoquine, termed the debrisoquine/sparteine polymorphism, is due to the absence of the debrisoquine 4-hydroxylase CYP2D1 protein (Gonzalez et al., 1989).

A cDNA coding for a male-specific testosterone 16α-hydroxylase was isolated from the mouse (Harada & Negishi, 1984; Wong et al., 1987), and was identified as being a member of the CYP2D gene subfamily. Comparison of the 16α-hydroxylase cDNA-deduced amino acid sequence with those of the rat CYP2D1 to CYP2D4 transcripts (Wong et al., 1987), revealed 82%, 72%, 78% and 70% sequence similarities, respectively.

2.4.2.1.5 The CYP2E gene subfamily

The metabolism of ethanol, acetone, acetoacetate, acetal and N-nitrosodimethylamine by a unique P450 enzyme has been observed in rats (Koop et al., 1982; Casazza et al., 1984; Koop & Casazza, 1985; Patten et al., 1986; Thomas et al., 1987; Wrighton et al., 1987). This unique P450 has been purified from rats (Ryan et al., 1985; Patten et al., 1986; Favreau et al., 1987), rabbits (Koop et al., 1982; Koop & Coon, 1984), and humans (Wrighton et al., 1987). The demethylation of N-nitrosodimethylamine, a compound that may play a role in nitrosamine-induced cancer in humans, is catalysed by a P450 belonging to the CYP2E gene subfamily found in humans (Wrighton et al., 1987; Umeno et al., 1988b). A single cDNA transcript, designated CYP2E1, has been identified in both
the rat and human (Song et al., 1986; Umeno et al., 1988a & b), as well as two in the rabbit (Khani et al., 1987 & 1988). The two CYP2E1 genes found in the rabbit share a 97% cDNA-deduced amino acid sequence similarity pointing towards a divergence from a common ancestor approximately 10 million years ago.

2.4.2.2 Regulation of the CYP2 gene family

Phenobarbitone (PB) has been reported to induce P450 enzymes in a variety of species ranging from bacteria to mammals, and it was therefore thought that the P450 induction mechanism would be conserved throughout (Waxman & Azaroff, 1992). cDNAs belonging to the CYP2 gene subfamily were among the first to be isolated and completely sequenced due to the fact that they are strongly induced by PB. Attempts to isolate and characterize the cytosolic receptor for PB has been unsuccessful and subsequently alternative mechanism(s) have been examined.

PB has been found to induce CYP2A1, CYP2B1/2, CYP2C6, CYP2C7, CYP2C11 in the rat, CYP2B4 in the rabbit, CYP2H1/2 in the chicken, and CYP102/106 in B. megaterium (Waxman & Azaroff, 1992). Although negative repressor-mediated regulation of the CYP2 genes occurs in prokaryotes, a relationship exists between PB-regulated prokaryotic and eukaryotic genes. This can be attributed to CYP2B1/2 gene induction. Here, rat CYP2B1/2 PB-mediated mRNA induction is blocked by cycloheximide (Bhat et al., 1987), whereas cycloheximide superinduced PB-mediated induction of chicken CYP2H1/2 mRNA (Hamilton et al., 1988), but inhibited CYP102/106 induction in B. megaterium (Fulco et al., 1991). Such observations may point towards variable drug-related interactions occurring in different species, particularly when looking at cycloheximide and PB.

PB regulatory elements (PBREs) have been identified in CYP2B1/2 and CYP102. These PBREs include a transcription initiation site (TIS) located 30bp upstream from the translation start site in CYP2B1/2, a modified TATA box found 20bp upstream of TIS (Atchison & Adesnik, 1983), and a 17bp Barbie Box at -227nt in CYP102 (He & Fulco, 1991). Other PB-inducible genes such as CYP2B1/2, α-acid glycoprotein, GSTs and CYP3A4 have also been identified as having a similar Barbie Box element (Liang et al., 1995). The sequence of such a Barbie Box was found within the cis-acting positive PBREs located at -98 to -69nt of the rat CYP2B1/2 genes (Upadhya et al., 1992). Three functionally important elements were identified in the CYP2B2 gene. These include FT1 (-36 to -8nt) and FT3 (-129 to -116nt), required for basal in vitro transcription, and FT2 (-66 to -42nt), needed for basal as well as PB-induced expression (Ramana & Kohli, 1998).

Distal regulatory sequences (DRS) identified in the CYP2B2 gene conferred PB-responsiveness on a heterologous promoter when spliced upstream of the promoter of a chloramphenicol acetyltransferase gene. The DRS identified contained an NF-1 binding site (Trottier et al., 1995). Studies performed suggest that the DRS plays a major role in CYP2B2 gene induction in response to PB, in contrast to proximal regulatory sequences (PRS), thereby questioning the role of the Barbie Box in the basal or PB-induced transcription of the CYP2B gene subfamily. One such study monitored the in vivo
expression of luciferase following injection of DRS and PRS-luciferase chimeric genes (lcg) directly into the rat liver (Park et al., 1996). The expression of the reporter gene was proportional to the increased copies of PRS-lcg, yet orientation independent. Here, the DRS-lcg (-2318 to -2155nt) was PB responsive, while the PRS-lcg (-110 to -1nt) was not (Park et al., 1996). A similar 177bp enhancer sequence required for PB-responsiveness, analogous to the rat CYP2B2 enhancer, was cloned into mice from the PB-responsive/non-responsive cyp2b10 and cyp2b9 genes, respectively. The DRS belonging to the cyp2b10 gene conferred PB-inducibility on a thymidine kinase promoter (Honkakoshi & Negishi, 1997). Footprint analysis revealed binding sites for NF-1 and NR in cyp2b10 but not in cyp2b9. The presence of putative binding sites for various nuclear factors have been identified within the rat CYP2B2 gene. These include sites for NF-1 and NR in cyp2b10 but not in cyp2b9. The presence of putative binding sites for various nuclear factors have been identified within the rat CYP2B2 gene. These include sites for NF-1 (-2188 to -2202nt), NF-xB (-3807 to -3798 and -879 to -870nt), HNF-3 (-3679 to -3669nt, -3174 to -3164nt and -2613 to -2603nt), and AP1 (-1428 to -1420nt) [Ramana & Kohli, 1998].

Sex-specific expression of the cyp2d9 gene can be seen in the domestic laboratory mouse, Mus musculus domesticus. A CpG site designated SD1 (sex difference information), located in a cis-acting element upstream of the cyp2d9 gene, is preferentially demethylated in males (Yoshioka et al., 1991). This preferential demethylation is also shared by another gene encoding a duplicate of compliment C4, designated slp (Yokomori et al., 1995). The natural C4 gene is expressed in both sexes with the CpG site becoming demethylated in both male and female mice, although, the slp gene SD1 is demethylated in males. The DNA binding factors GABPa/β heterodimer and NF2d9 interact with the SD1 sequence and are sensitive to the CpG (-97nt) and T/C base change (-99nt) within the SD1 gene (Gonzalez & Lee, 1996). Here, GABP acts as a preferential transcriptional activator for the male-specific cyp2d9 promoter, while the functional role of the NF2d9 protein with respect to the latter promoter, remains to be determined. The role of GABP and/or NF2d9 in the male-specific expression of cyp2d9 in M. domesticus is still unknown (Gonzalez & Lee, 1996).

A model describing the induction mechanism of the rat CYP2B1/2 genes by PB encompassed the preferential binding of a dephosphorylated 26kDa protein to negative PBRE (-160 to -126nt) [Prabhu et al., 1995; Sultana et al., 1997]. The uninduced dephosphorylated 26kDa protein forms a complex with a heme-binding 65kDa protein and a 94kDa protein (Sultana et al., 1997), and preferentially binds positive rather than negative PBRE. The protein complex does not interact with DRS and represents the basal expression of CYP2B2. Interaction of the complex, when tethered to positive PBRE, with proteins bound to DRS (-2318 to -2155nt) is the switch, and is catalyzed by the phosphorylation of the 26kDa protein by PB (Fig. 2.3). The binding of the phosphorylated protein complex to the DRS elicits the maximum transcriptional activation of the CYP2B2 gene by PB (Prabhu et al., 1995). cAMP-dependent protein kinase (PKA) may be the enzyme responsible for the phosphorylation of the transcription factors due to the fact that PB has been found to stimulate acidic nuclear protein phosphorylation in the rat liver (Blankenship & Bresnick, 1974). In vivo activation of cAMP, PKA and RNA polymerase by PB occurs prior to the induction of drug metabolizing enzymes (DME) in the rat liver (Byus et al., 1976; Manen et al., 1978).
The reports suggest that the phosphorylation of transcription factors by specific protein kinases form part of pretranscriptional processes involved in the regulation of PB-enhanced transcription of CYP2B genes, and therefore support the model.

The developmental activation of DBP, a protein factor controlling adult liver-specific expression of the rat albumin gene (Mueller et al., 1990), correlates with the transcriptional activation of the rat CYP2D6 gene at the onset of puberty. While the rat albumin gene is bound and activated by C/EBPα, C/EBPβ, and DBP, CYP2D6 can bind all three factors but is only activated by DBP. The CYP2D6 also contains a functional binding site for the nuclear receptor HNF-3, and can be induced by PB via a PBRE (Friedberg et al., 1986). The PBRE remains to be identified. Functional binding sites for HNF-4 can also be found in the rabbit CYP2C1, CYP2C2, and CYP2C3 genes (Venepally et al., 1992; Chen et al., 1994).

Upon examination of the five CYP2D genes present in the rat, designated CYP2D1 to CYP2D5 (Matsunaga et al., 1989), differences in developmental activation of CYP2D3 and CYP2D5 can be attributed to the methylation of certain cytosine residues (Gonzalez & Lee, 1996). The stimulation of the promoter region involved in trans-activation of the CYP2D5 gene requires the enhancer-binding protein C/EBPβ, as well as the ubiquitous factor Sp1 (Gonzalez & Lee, 1996). Sp1 alone promotes a low level of CYP2D3 transcription while C/EBPβ in unable to bind its recognition sequence in the absence of Sp1 (Fig. 2.4). The presence of Sp1, C/EBPβ, TAF110 (the TBP factor), and another unconfirmed TAF, promote higher levels of CYP2D5 transcription (Gonzalez & Lee, 1996).

Figure 2.3. Gene regulation of the Cytochromes P4502B by Barbiturates (Ramana & Kohli, 1998).

The developmental activation of DBP, a protein factor controlling adult liver-specific expression of the rat albumin gene (Mueller et al., 1990), correlates with the transcriptional activation of the rat CYP2D6 gene at the onset of puberty. While the rat albumin gene is bound and activated by C/EBPα, C/EBPβ, and DBP, CYP2D6 can bind all three factors but is only activated by DBP. The CYP2D6 also contains a functional binding site for the nuclear receptor HNF-3, and can be induced by PB via a PBRE (Friedberg et al., 1986). The PBRE remains to be identified. Functional binding sites for HNF-4 can also be found in the rabbit CYP2C1, CYP2C2, and CYP2C3 genes (Venepally et al., 1992; Chen et al., 1994).

Upon examination of the five CYP2D genes present in the rat, designated CYP2D1 to CYP2D5 (Matsunaga et al., 1989), differences in developmental activation of CYP2D3 and CYP2D5 can be attributed to the methylation of certain cytosine residues (Gonzalez & Lee, 1996). The stimulation of the promoter region involved in trans-activation of the CYP2D5 gene requires the enhancer-binding protein C/EBPβ, as well as the ubiquitous factor Sp1 (Gonzalez & Lee, 1996). Sp1 alone promotes a low level of CYP2D3 transcription while C/EBPβ in unable to bind its recognition sequence in the absence of Sp1 (Fig. 2.4). The presence of Sp1, C/EBPβ, TAF110 (the TBP factor), and another unconfirmed TAF, promote higher levels of CYP2D5 transcription (Gonzalez & Lee, 1996).
The regulation of CYP2E1 by xenobiotics such as ethanol, acetone, chloroform, and nitrosamines differs from other P450s in that certain patho-physiological conditions such as a high fat diet, diabetes, and fasting readily induce the gene (Song et al., 1986; Johansson et al., 1988). The administration of ethanol increases CYP2E1 protein synthesis without an increase in CYP2E1 mRNA. This points towards post-transcriptional stabilization of the CYP2E1 protein by a xenobiotic. The modification of CYP2E1 by both acetone and ethanol alters the proteins conformation, rendering it resistant to proteolytic degradation (Eliasson et al., 1990 & 1992; Roberts, 1997; Song et al., 1989). CYP2E1 mRNA stabilization was also observed in spontaneously diabetic rats. This might point towards a mechanism or signal sequence that specifically governs CYP2E1 mRNA stabilization relative to other liver mRNA (Favreau et al., 1987; Dong et al., 1988).

2.4.3 Glucocorticoid-inducible Cytochrome P450 monooxygenases

2.4.3.1 The P450III gene family

The synthetic steroid pregnenolone 16α-carbonitrile (PCN), induces two P450s from the P450III (CYP3) gene subfamily. The first, designated PCN1/CYP3A1 (Gonzalez et al.,
1986), was purified from the rat (Elshourbagy & Guzelian, 1980), along with the second male-specific CYP3 P450 purified (Guengerich et al., 1982; Waxman, 1984), designated PCN2/CYP3A2 (Gonzalez et al., 1986). Immunoinhibition, immunoquantitation, and lipid reconstitution studies have established that the PCN-induced P450 enzymes belonging to the CYP3 gene subfamily have catalytic activities towards testosterone, ethylmorphine N-demethylation, and extracted microsomal lipids (Elshourbagy & Guzelian, 1980; Waxman et al., 1985; Nagata et al., 1987; Yamazoe et al., 1988). A human P450 enzyme, designated P450NF, mainly responsible for inhibiting nifedipine metabolism, has been found to be related to the male-specific rat CYP2C12 P450 enzyme, also involved in nifedipine oxidation (Guengerich et al., 1986).

Both CYP3A1 and CYP3A2 from the rat share 90% nucleotide sequence similarity (Gonzalez, 1989). Of three CYP3 genes identified in the rabbit, P4503c was the first to be isolated (Gonzalez, 1989), along with three CYP3 genes present in man, namely HLP (Molowa et al., 1986), P450NF (Guengerich et al., 1986; Beaune et al., 1986) and hPCN3 (Aoyama et al., 1989).

2.4.3.2 Regulation of the CYP3 gene family

The expression of drug-metabolizing enzymes is enhanced through transcriptional and post-transcriptional mechanisms by naturally occurring steroids, as well as their synthetic analogues. The synthetic glucocorticoid pregnenolone 16a-carbonitrile (PCN) and dexamethasone (DEX), form part of this group. Enzymes such as prolactin, UDP-glucuronyl transferase, the epoxide hydrolases, θ-glutamyl transpeptidase, NADPH cytochrome P450 oxidoreductase, tyrosine amino transferase, and the corresponding CYP2E1 and CYP3A1 proteins, occur in a variety of species and are all regulated by steroids (Simmons et al., 1987).

Regulation of the rat CYP3A1 and CYP3A2 genes is restricted by the animals age and sex, as well as the type of inducer employed. CYP3A1 mRNA is absent in untreated cells and therefore relies on inducers to regulate its expression (Gonzalez et al., 1986). The CYP3A1 gene is also transcriptionally activated by glucocorticoids (Simmons et al., 1987). On the other hand, CYP3A2 is not inducer-dependent as it is constitutively expressed in adult male rats, while being transiently expressed in young females (Gonzalez et al., 1986).

Work done by Quattrochi et al. (1995) investigated glucocorticoid CYP3A gene activation. In an uninduced cells cytosol, hsp70 and hsp90 are associated with glucocorticoid receptor (GR). Interaction between glucocorticoid and the GC complex results in dissociation of hsp70. Here, hsp70 served to facilitate the formation of the complex between GR and hsp90 (Ramana & Kohli, 1998). The formation of a binary complex, comprising GR and glucocorticoid, result in the dissociation of hsp90. The binary complex is then translocated to the nucleus, where it binds glucocorticoid responsive element (GRE) as well as other proteins. Proteins already bound to PRS, located between -148 and -71nt of CYP3A1, contain two responsive elements named FP1 and FP2. The FP1 sequence, situated between -148 and -115nt, is both DEX and PCN
responsive. Quattrochi et al. (1995) proposed that CYP3A1 gene transcription involved the specific binding of proteins to this fragment, finally activating transcription (Fig. 2.5).

![Diagram of gene regulation of the Cytochromes P4503A by Glucocorticoids (Ramana & Kohli, 1998).](image)

Footprint analysis of the PRS (-165 to -73nt) of the CYP3A2 gene identified a nuclear protein binding site, designated 6βA-A (-106 to -87nt). This nuclear protein binding site was found to be specific for binding hepatocyte nuclear factor-4 (HNF-4), presumed to be needed for CYP3A2 basal transcription (Ramana & Kohli, 1998).

Post-translational substrate-dependent stabilization events also play a role in the regulation of the CYP3A genes in the rabbit and the rat. This has been reported for troleandomycin (TAO) and erythromycin, a class of antibiotics. Administration of TAO caused an increase in CYP3A1/CYP3A2 mRNA levels in rats (Watkins et al., 1986; Gonzalez et al., 1986), while both TAO and erythromycin increased the half-life of the CYP3A1 isozyme in cultured rat hepatocytes (Watkins et al., 1986). On the other hand, DEX increased mRNA levels in in vitro cultures and rats, suggesting that DEX and TAO induce CYP3 proteins via different mechanisms. These mechanisms constitute DEX induction through transcriptional activation (Simmons et al., 1987), and TAO via mRNA stabilization (Watkins et al., 1986).

Trans-activation and in vitro DNA-binding studies demonstrated that the CYP3A10 gene is preferentially expressed in male hamsters, and contains a cis-binding element capable of binding activated signal transduction/activator of transcription (STAT) proteins (Fig. 2.6) [Subramanian et al., 1995]. Activation of the liver STAT5-related factor occurs within minutes of growth hormone treatment while male-specific expression of the
CYP2C12 gene in rats takes 24-48 hours (Waxman et al., 1995). As a result, it is uncertain as to whether the growth hormone-mediated sex-dependent expression of P450 genes is due to the stimulation of STAT proteins by the growth hormone receptor (Gonzalez & Lee, 1996).

2.4.4 Peroxisome Proliferator-inducible Cytochrome P450 monooxygenases

2.4.4.1 The P450IV gene family

Hypolipidemic drugs such as nafenopin, halogenated aromatic solvents, ciprofibrate, and clofibrate are known to induce ω-hydroxylases and to proliferate peroxisomes in several species (Lock et al., 1989). A clofibrate-inducible P450 able to metabolize lauric acid at its ω position, designated CYP4A1 (Tamburini et al., 1984), was isolated and sequenced from the rat and represents the first member of the CYP4A gene family (Hardwick et al., 1983). CYP4A1 is also induced by other hypolipidemic agents, has no detectable activity towards testosterone, benzphetamine and ethoxyresorufin, but is active in the oxidation of arachidonic acid (Bains et al., 1985).

Included in the CYP4A gene family are the cDNAs designated CYP4A2 and CYP4A3. CYP4A3 has been found to have a 72% cDNA-deduced amino acid sequence similarity with that of the CYP4A1 gene (Gonzalez, 1989). A P450 expressed in the lungs of pregnant rabbits has been purified and isolated (Williams et al., 1984; Matsubara et al.,
1987), and is designated CYP4A6. Both the CYP4A1 and CYP4A2 genes were found to possess 13 and 12 exons, respectively (Gonzalez, 1989). Based upon the assumption that exon loss has occurred during the evolution of the eukaryotic genome (Gilbert et al., 1986), and the fact that the CYP4A gene family has considerably more exons when compared to the CYP1 and CYP2 gene families, it seems likely that the CYP4A gene family evolved early in the evolution of the P450 gene superfamily (Degtyarenko & Archakov, 1993).

2.4.4.2 Regulation of the CYP4 gene family

The CYP4AJ and CYP4A2 genes form part of a gene battery including the enoyl CoA hydratase, 3-hydroxyacyl CoA dehydrogenase, and fatty acyl CoA oxidase (ACO) genes. The transcription rate of this gene battery is increased upon administration of clofibrate, a hypolipidemic drug, which in turn proliferates peroxisomes in various species (Lock et al., 1989), and as a result are termed peroxisomal proliferators (Ramana & Kohli, 1998).

Significant progress has been made in understanding the mechanisms involved in the induction of the CYP4A gene family by peroxisome proliferators. The molecular induction mechanisms involved might be mediated by a member of the nuclear receptor supergene family of ligand-activated transcription factors, viz, the peroxisomal proliferator activated receptor (PPAR) [Green, 1992; Poellinger et al., 1992]. Three different PPARs have been identified (α, β/δ, and γ), and each displays a distinct expression pattern (Schoonjans et al., 1996). Whether PP binds PPAR or acts via second messengers remains to be established (Kliewer et al., 1992). CYP4A gene expression is mediated by the formation of a binary complex between retinoid X receptor (RXR) in the nucleus, and PPAR. This heterodimer (PPAR-RXR) then binds PP responsive element (PPRE) at the 5'-flanking regions of the CYP4A gene, thereby enhancing gene expression (Fig. 2.7) [Ramana & Kohli, 1998].

![Diagram of gene regulation of the Cytochromes P4504A by Peroxisome Proliferators](image-url)

**Figure 2.7.** Gene regulation of the Cytochromes P4504A by Peroxisome Proliferators (Ramana & Kohli, 1998).
Endogenous fatty acids such as docosahexaenoic acid (22:6), arachidonic acid (20:4), linolenic acid (18:3), and lauric acid (12:0), as well as sulphur-substituted fatty acids such as tetradecylthioacetic acid and tetradecylthiopropionic acid, act as pre-translational regulators of CYP4A1 (Tolet et al., 1994). Kroetz et al. (1998) showed that CYP4A1, CYP4A2 and CYP4A3 are induced via insulin-dependent diabetes and starvation, thereby demonstrating that the requisite cellular changes necessary for the signaling of a member of a nuclear receptor family (PPARα), can be produced in a patho-physiological state (Kroetz et al., 1998). Xenobiotic-mediated CYP2A, CYP3A and CYP4A gene expression is also suppressed by low insulin levels in cultured primary rat hepatocytes (Woodcroft & Novak, 1999), illustrating that insulin can have multiple and varied effects on gene expression.

Characterization of an 8kb 5'-flanking region of DNA from CYP4A6 has revealed several cognitive sequences of various transcription factors (Muerhoff et al., 1992b). These transcription factors constitute HNF-1 and HNF-4 at -20 to -8nt and -106 to -4nt, respectively, a consensus Sp1 recognition sequence at -46bp of TIS, and a PP-inducible PPAR-dependent DRS (-782 to -703nt). Of the three PPREs located in the CYP4A6 gene within the DRS and PRS, the PPRE located at -667 to -644nt is the most responsive (Palmer et al., 1994 & 1995; Ramana & Kohli, 1998).

In summary, extracellular and intracellular signals are transformed into cellular responses via the triggering of nuclear receptor (NR) target genes by various transcription factors. The NR superfamily, which codes for the various transcription factors, shares significant similarity with the thyroid hormone and classical steroid-hormone receptors in their ligand binding domains (LBDs), as well as their DNA-binding domains (DBDs) [Lund et al., 1991; Gronemeyer, 1992; Reichel & Jacob, 1993; Kemper, 1998]. The NR superfamily also contains receptors for non-steroid ligands including prostaglandins, retinoid acid as well as fatty acids (Mangelsdorf et al., 1995).

Xenobiotics enhance the expression of the CYP genes through receptor and nonreceptor mediated mechanisms. PAH, glucocorticoids and PP enhance CYP1A, 3A and 4A expression via receptor mediated mechanisms while barbituates do so by means of nonreceptor mediated mechanisms. NRs responsible for the regulation of various CYP genes are activated by ligands that have in turn been degraded by several CYP genes. Examples of such feedback regulation can be seen in the CYPs metabolizing retinoids [CYP2C7 and the retinoic acid receptor (RAR)], PPARα-activated CYP4A, as well as CYP3A and CYP2B. Here, steroid hormones and fatty acids act as ligands for the pregnane X receptor (PXR) and PPARα, which in turn increase ligand elimination by CYP3A-catalyzed 6β-hydroxylation and CYP4A-mediated ω-hydroxylation, respectively (Gronemeyer, 1992; Reichel & Jacob, 1993; Lund et al., 1998; Ramana & Kohli, 1998). Additional candidates which could be added to this class might include the stereospecific androgen 7α- and 16α-hydroxylase CYP2A isoforms (Waxman et al., 1983; Wong et al., 1987; Gonzalez, 1989). On the other hand, some CYPs ensure the elimination of a wide range of endo- or xenobiotics. Here, basal and liver-specific CYP gene expression is controlled by orphan receptors such as HNF-4. CYP genes that belong to this group have
no activating ligand or feedback cycle and include CYP2A, CYP2C, CYP2D and CYP3A (Ramana & Kohli, 1998; Gonzalez & Lee, 1996).

2.5. Fungal Cytochrome P450 monoxygenases

Unlike their mammalian counterparts, the fungal cytochrome P450 enzyme systems have only received attention over the last decade. Low specific content and sensitivity to homogenization and solubilization have hampered studies of these fungal enzymes (Sariaslani, 1991). This being the case, recent advances in molecular techniques have promoted the isolation and characterization of cytochrome P450 monoxygenases in both yeast and filamentous fungi.

Fungal cytochrome P450s have been shown to be important in both biosynthetic and catabolic reactions. The roles played by yeast and filamentous fungi in the oxidation of aromatic compounds as well as the hydroxylation of alkanes, fatty acids and drugs, have been investigated (Kapoor & Lin, 1984; Fukui & Tanaka, 1981; Mauersberger et al., 1996). Consequently, the occurrence of cytochrome P450 enzymes in various fungi, and their subsequent roles in aromatic hydrocarbon degradation, have been described. Some of the fungi investigated thus far include the white rot fungus Phanerochaete chrysosporium, Fusarium oxysporum, Fusarium sporotrichioides, Rhodospirillum toruloides, Aspergillus nidulans, Nectria haematococca, and Rhodotorula minuta (Kullman & Matsumura, 1997; Hohn et al., 1995; Kizawa et al., 1991; Fujii et al., 1989; Kelkar et al., 1997; Maloney & VanEtten, 1994).

Both R. toruloides and R. minuta contain benzoate-inducible cytochrome P450 monoxygenases, although increased levels of cytochrome have been correlated with isobutene-forming activities for R. minuta (Muncnerová & Augustin, 1995; Fujii et al., 1989; Fabián & Degtyarenko, 1997). The R. minuta P450 has been sequenced and is designated CYP35B1 (Fabián & Degtyarenko, 1997). The A. nidulans cytochrome P450 monoxygenase gene stcl/CYP39 is located within a cluster of genes required for the conversion of versicolorin A to sterigmatocystin (ST) [Keiler et al., 1995; Kelkar et al., 1997]. Here, aflatoxin and ST are carcinogenic polyketides produced by several Aspergillus spp. which, when found in food and feed supplies, could present serious health and economic threats to the public. Here A. nidulans utilizes a cytochrome P450 monoxygenase for the production of aflatoxin and ST, while the fungal plant pathogen N. haematococca has been found to convert pisatin, an isoflavonoid phytoalexin, to a less toxic derivative via a cytochrome P450 monoxygenase-catalyzed demethylation reaction (Miao et al., 1991; Maloney & VanEtten, 1994). As is seen in A. nidulans, the gene encoding the cytochrome P450 in N. haematococca, designated PDAT9, forms part of a group of genes (PDA genes) used to identify different phenotypes of pisatin demethylating activity (Maloney & VanEtten, 1994). Only two of the three existing phenotypes can cause disease on pea, implying that rapid and efficient detoxification of this plant antibiotic is required for pathogenicity. The cytochrome P450 monoxygenase genes CYP55A1 and CYP58 found in F. sporotrichioides and F. oxysporum are involved in the reduction of nitrite and trichothecene biosynthesis, respectively (Kizawa et al., 1991; Hohn et al., 1995). Similarly, as is seen for A. nidulans and N. haematococca, the
F. sporotrichioides cytochrome P450 gene (CYP55A1) forms part of at least three genes closely involved in trichothecene biosynthesis (Hohn et al., 1995). Although a novel cytochrome P450 monooxygenase gene (CYP63-IA) has been identified in the white rot fungus P. chrysosporium, questions still remain regarding the role played by CYP63-IA in both fungal physiology and xenobiotic transformation (Kullman & Matsumura, 1997).

2.5.1. The CYP51 gene family involved in ergosterol biosynthesis

The biosynthesis of the membrane sterol ergosterol has been studied in both filamentous fungi and yeast. The production of ergosterol is dependent on the metabolic activity of the P45014 DM-encoding genes. The CYP51 gene product, sterol 14α-demethylase (P45014 DM), is expressed in all three eukaryotic phyla, and is believed to be one of the oldest cytochrome P450 monooxygenase genes. Although the CYP51 genes present in the three phyla are believed to have a common ancestor, a closer analysis of the fungal, vertebrate and plant CYP gene structures could lead the identification of the closest CYP51 gene relatives (Rozman et al., 1996).

P45014 DM is involved in the biosynthesis of ergosterol in fungi and does so by removing the C32-methyl group at the 14α-position from precursor sterols. Antifungal agents are able to block P45014 DM activity and are subsequently used in the treatment of immunocompromised patients suffering from various fungal infections (Venkateswarlu et al., 1995; Lamb et al., 1995). The inhibitory effect of antifungal agents towards lanosterol 14α-demethylase enabled Kalb et al. (1986) to exploit this property thereby isolating the lanosterol 14α-demethylase gene, ERG11, from S. cerevisiae. It was discovered that ERG11 mRNA levels were affected by glucose, the presence of heme, as well as the anaerobic growth rate of the culture (Turi & Loper, 1992). Additional information obtained regarding multiple regulatory elements controlling ERG11 gene expression include: the activation of an upstream activating sequence (UAS1) by the HAP1 protein product; the involvement of the ROX1 and Old (overexpression of lanosterol demethylase) repressors in the repression of ERG11 expression; as well as the coordinated regulation of NADPH-cytochrome P450 reductase (CPR1) with ERG11 (Turi & Loper, 1992). The information obtained indicates that ERG11 should become a member of the hypoxic gene family found to include HEM13, CYC7, COX3b, and ANB1. Additional yeast CYP51 genes have been isolated from Candida tropicalis, Candida albicans, Candida glabrata, Candida guilliermondii, Candida kefyr, Candida parapsilosis and Schizosaccharomyces pombe (Kalb et al., 1986; Chen et al., 1987; Lai & Kirsch, 1989; Morace et al., 1991, Fábián & Degtyarenko, 1997). With the exception of the Penicillium italicum CYP51 gene, the remaining yeast CYP51 genes identified thus far are intronless (Rozman et al., 1996).

2.5.2. The CYP52 gene family involved in alkane utilization

Apart from the fungal P450s involved in biosynthesis and detoxification, there exist yeasts involved in alkane assimilation that exhibit different metabolic features when compared to those utilized for metabolic and physiological studies. The utilization of n-alkanes as a sole carbon and energy source by yeasts belonging to the genus Candida
have demanded considerable attention. These yeasts have inducible cytochrome P450
monooxygenases that are only detected when grown on n-alkanes. An additional
characteristic exhibited by these yeasts is the proliferation of peroxisomes in the presence
of n-alkanes. The characterization of n-alkane-inducible monooxygenase genes have been
reported for *C. tropicalis*, *C. maltosa*, *C. apicola* and *Y. lipolytica* (Seghezzi et al., 1992;
Zimmer et al., 1996; Lottermoser et al., 1996; Iida et al., 1998 & 2000). Each yeast
species was found to contain several cytochrome P450 monooxygenase genes involved in
n-alkane assimilation. As a result, the individual families of genes were assigned to a new
P450 gene family called the yeast CYP32 multigene family (Nebert et al., 1991).

The pathway and enzymes involved in alkane assimilation by *C. maltosa* and *C.
tropicalis* have been studied in detail. Two different hydroxylating reactions have been
attributed to alkane-inducible P450 systems. The first system constitutes the terminal
hydroxylation of n-alkanes initiating the degradation of these growth substrates, and the
second, the ω-hydroxylation of fatty acids (Fukui & Tanaka, 1981; Porter & Coon, 1991;
Gonzalez, 1989; Sanglard & Loper, 1989). Investigations showed that intermediates with
a chain length equivalent to that of the substrate are short lived. The large amounts of C16
and C18 fatty acids produced indicate even-numbered fatty acid chains shortened by β-
oxidation. The action of cerulenin in inhibiting de novo synthesis of fatty acids (Tanaka
et al., 1976), suggests that de novo fatty acid synthesis in *C. maltosa* grown on alkanes is
negligible, even though cellular fatty acid concentrations are significantly large (Blasig
et al., 1989). The terminal hydroxylation activities of *C. tropicalis* toward two model
substrates, lauric acid and hexadecane respectively, have been shown to result from
specific proteins. These proteins, coded for by various alkane-inducible genes, have
defined substrate specificities (Seghezzi et al., 1992).

The hydroxylation of oleic acid by cell-free extracts of *C. bombicola* was the first report
suggesting the involvement of a P450 in sophorose lipid production (Heinz et al., 1969 &
1970). The enzyme system was classified as a mixed-function oxidase that was inhibited
by CO, suggesting that a microsomal cytochrome P450 participated as the O2-activating
component. Work done later by Hommel et al. (1994), showed an increased cytochrome
P450 and cytochrome oxidase aa3 content in both the logarithmic and stationary growth
phases of *C. apicola*, another sophorose lipid producing yeast. The cytochrome P450
content was therefore associated with the initiation of sophorose lipid biosynthesis
(Hommel et al., 1994). The production of sophorose lipids by *C. apicola* varies in fatty
acid composition as it is capable of utilizing many long-chain aliphatic compounds
(Jones, 1968; Jones & Howe, 1968a, b & c). The composition of sophorose lipids
therefore depends on the substrate added. Both long-chain and short-chain fatty acid
precursors are poorly incorporated, yielding low concentrations of sophorose lipids. C10
to C19 alkanes, fatty acids and derivatives thereof, are directly converted to [ω-1]- or ω-
hydroxy fatty acids prior to incorporation into sophorose lipids. As a result, the chain
length of the fatty acid substrate determines the hydroxylation position. The detection of
cytochrome P450s in cells in the stationary growth phase following growth on either
glucose, hexadecane, octadecane or mixtures thereof (Jones & Howe, 1968a; Weber et
al., 1990), suggested that hydroxylation reactions were involved and that these reactions
were catalyzed by P450s having different substrate- and regiospecificities (Hommel et al., 1990).

Similarly, *Y. lipolytica* had been known to assimilate *n*-alkanes, 1-alkenes and chlorinated alkanes as carbon sources, although no cytochrome P450 monooxygenases had at that time been identified (Nyns et al., 1967; Murphy & Perry, 1984). Five phenotypic classes of alkane-nonutilizing *Y. lipolytica* mutants (alkA to alkE) were isolated and characterized, depending on the availability of different alkane degradation pathway intermediates (Kujau et al., 1992; Heslot, 1990; Barth & Gaillardin, 1997). The uptake of alkanes was found to be inducible and reliant on active transport, although the mutations in 16 loci caused a significant reduction in *n*-alkane uptake. Subsequently, three subtypes of alkA mutants were characterized (Heslot, 1990; Barth & Gaillardin, 1997). These included alkAa mutants unable to utilize all types of *n*-alkanes, alkAb mutants responsible for C₈ to C₁₂ alkane assimilation, and alkAc mutants predominantly involved in long-chain (>C₁₂) *n*-alkane uptake (Heslot, 1990; Barth & Gaillardin, 1997). This data pointed towards the existence of either specific cytochrome P450 monooxygenases, or several length specific *n*-alkane-uptake systems present in *Y. lipolytica*.

To date 25 different *CYP52* genes have been sequenced from yeasts belonging to the genera *Yarrowia* and *Candida* (Mauersberger et al., 1984; Käppeli et al., 1986; Schunck et al., 1987; Hommel et al., 1994; Heslot, 1990). These genes, their gene products as well as the substrates regulating their expression are summarized in Table 2.1.
TABLE 2.1. A TABLE OF THE YEAST CYP52 GENES, THEIR GENE PRODUCTS AND THE SUBSTRATES THEY ARE INDUCED AND REPRESSED BY.

<table>
<thead>
<tr>
<th>Yeast</th>
<th>Gene Symbol</th>
<th>Trivial name</th>
<th>Inducing Substrates</th>
<th>Repressing Substrates</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. tropicalis</em></td>
<td>CYP52A1</td>
<td>Alk1</td>
<td>Alkanes + acids</td>
<td>Glucose</td>
<td>Sanglard &amp; Loper, 1989</td>
</tr>
<tr>
<td></td>
<td>CYP52A2</td>
<td>Alk2</td>
<td>Alkanes + acids</td>
<td>Glucose</td>
<td>Seghezzi et al., 1991</td>
</tr>
<tr>
<td></td>
<td>CYP52A6</td>
<td>Alk3</td>
<td>Alkanes + acids</td>
<td>Glucose</td>
<td>Seghezzi et al., 1992</td>
</tr>
<tr>
<td></td>
<td>CYP52A7</td>
<td>Alk4</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Seghezzi et al., 1992</td>
</tr>
<tr>
<td></td>
<td>CYP52A8</td>
<td>Alk5</td>
<td>Alkanes</td>
<td>Glucose</td>
<td>Seghezzi et al., 1992</td>
</tr>
<tr>
<td></td>
<td>CYP52B1</td>
<td>Alk6</td>
<td>Dodecanol</td>
<td>C_{10}-C_{12} alkanes</td>
<td>Seghezzi et al., 1992</td>
</tr>
<tr>
<td></td>
<td>CYP52C1</td>
<td>Alk7</td>
<td>C_{10}-C_{12} alkanes</td>
<td>Glucose</td>
<td>Seghezzi et al., 1992</td>
</tr>
<tr>
<td><em>C. maltosa</em></td>
<td>CYP52A3</td>
<td>Alk1A/Cm1</td>
<td>C_{12}-C_{16} alkanes</td>
<td>Glucose</td>
<td>Schunck et al., 1989</td>
</tr>
<tr>
<td></td>
<td>CYP52A4</td>
<td>Alk3A/Cm2</td>
<td>Acids + alkanes</td>
<td>Glucose</td>
<td>Schunck et al., 1991</td>
</tr>
<tr>
<td></td>
<td>CYP52A5</td>
<td>Alk2A/Cm3</td>
<td>Alkanes + acids</td>
<td>Glucose</td>
<td>Ohkuma et al., 1991b</td>
</tr>
<tr>
<td></td>
<td>CYP52A9</td>
<td>Alk5A</td>
<td>C_{12}-C_{16} acids</td>
<td>Glucose</td>
<td>Zimmer et al., 1996</td>
</tr>
<tr>
<td></td>
<td>CYP52A10</td>
<td>Alk7A</td>
<td>C_{12} acids</td>
<td>Glucose</td>
<td>Zimmer et al., 1998</td>
</tr>
<tr>
<td></td>
<td>CYP52A11</td>
<td>Alk8A</td>
<td>C_{12} acids</td>
<td>Glucose</td>
<td>Zimmer et al., 1998</td>
</tr>
<tr>
<td></td>
<td>CYP52C2</td>
<td>Alk6A</td>
<td>Acids + alcohols</td>
<td>Unknown</td>
<td>Ohkuma et al., 1998</td>
</tr>
<tr>
<td></td>
<td>CYP52D1</td>
<td>Alk4A</td>
<td>Acids + alcohols</td>
<td>Unknown</td>
<td>Ohkuma et al., 1998</td>
</tr>
<tr>
<td><em>C. apicola</em></td>
<td>CYP52E1</td>
<td>None</td>
<td>Long-chain acids</td>
<td>Unknown</td>
<td>Lottermoser et al., 1996</td>
</tr>
<tr>
<td></td>
<td>CYP52E2</td>
<td>None</td>
<td>Long-chain acids</td>
<td>Unknown</td>
<td>Lottermoser et al., 1996</td>
</tr>
<tr>
<td><em>Y. lipolytica</em></td>
<td>CYP52F1</td>
<td>Y/Alk1</td>
<td>Alkanes</td>
<td>Glycerol</td>
<td>Iida et al., 1998 &amp; 2000</td>
</tr>
<tr>
<td></td>
<td>CYP52F2</td>
<td>Y/Alk2</td>
<td>Hexadecane</td>
<td>Glycerol</td>
<td>Iida et al., 2000</td>
</tr>
<tr>
<td></td>
<td>CYP52G1</td>
<td>Y/Alk3</td>
<td>&gt;C_{16} alkanes</td>
<td>Glycerol</td>
<td>Iida et al., 2000</td>
</tr>
<tr>
<td></td>
<td>CYP52H1</td>
<td>Y/Alk4</td>
<td>&gt;C_{16} alkanes</td>
<td>Glycerol</td>
<td>Iida et al., 2000</td>
</tr>
<tr>
<td></td>
<td>CYP52I1</td>
<td>Y/Alk5</td>
<td>&gt;C_{16} alkanes</td>
<td>Glycerol</td>
<td>Iida et al., 2000</td>
</tr>
<tr>
<td></td>
<td>CYP52J2</td>
<td>Y/Alk6</td>
<td>&gt;C_{16} alkanes</td>
<td>Glycerol</td>
<td>Iida et al., 2000</td>
</tr>
<tr>
<td></td>
<td>CYP52J3</td>
<td>Y/Alk7</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Iida et al., 2000</td>
</tr>
<tr>
<td></td>
<td>CYP52J1</td>
<td>Y/Alk8</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Iida et al., 2000</td>
</tr>
</tbody>
</table>

Northern blot analysis showed that six of the seven *C. tropicalis* CYP52 genes are induced by different aliphatic carbon sources. All are repressed by glucose, with the exception of CYP52B1 and CYP52A7. Here, CYP52B1 produced a weak but unambiguous signal in the presence of glucose whilst CYP52A7 produced no mRNA transcripts under the conditions tested (Seghezzi et al., 1992). Glucose repression is also witnessed for seven of the eight *C. maltosa* CYP52A genes (Zimmer et al., 1996 & 1998; Ohkuma et al., 1998). In general, the *C. tropicalis* CYP52A1, A2, and A6 genes showed stronger induction when compared to CYP52A7, A8, B1 and C1 on all the substrates tested, particularly the substituted aliphatic substrates such as oleic acid and lauric aldehyde (Seghezzi et al., 1992). CYP52A2 showed its strongest induction by n-alkanes. Subsequent heterologous expression of the *C. tropicalis* CYP52 genes in *S. cerevisiae* showed CYP52A6 to have a strong hydroxylating activity towards hexadecane. Initially, induction experiments showed CYP52A7 to be a pseudogene (Seghezzi et al., 1992).
however, successful expression of the gene in *S. cerevisiae* refuted this initial assumption. Although *CYP52A8* appeared to be weakly induced by substituted aliphatic substrates, spectral and immunological analysis showed a clear substrate preference for lauric acid (Seghezzi et al., 1992).

The functional n-alkane hydroxylating efficiency of the four major *C. maltosa* P450<sub>alk</sub> isoforms, *CYP52A3, A4, A5, and A9* investigated by Zimmer et al. (1996), showed *CYP52A3* to have a distinct alkane hydroxylating ability. *In vivo* reconstitution of *CYP52A4* and *A3* exhibited relatively broad substrate specificities, while *CYP52A9* preferentially catalyzed the hydroxylation of fatty acids, showing a clear chain-length dependency. *CYP52A3* was shown to be the most versatile of the four isoforms when considering its enzymatic activity and ability to sustain growth on different chain-length n-alkanes (Ohkuma et al., 1998). Both *CYP52A10* and *A11* are unable to initiate alkane degradation, although their short-chain fatty acid hydroxylation ability may be important for the production of dicarboxylic acids destined for the peroxisomal β-oxidation system (Zimmer et al., 1998). Both *CYP52C2* and *D1* have been observed to be induced at low levels by n-alkanes, n-alcohols, and fatty acids (Ohkuma et al., 1998).

The first cytochrome P450s associated with the sophorose lipid-producing yeast *C. apicola* were isolated and characterized by way of a heterologous primer PCR-mediated approach (Lottermoser et al., 1996). Overall global alignments of the *C. apicola* *CYP52E1* and *CYP52E2* genes with other P450s showed the highest overall amino acid sequence similarity with the mammalian *CYP4A* subfamily as well as the *B. megaterium* P450, P450<sub>Bm-3</sub> (Lottermoser et al., 1996). Here, the *CYP4A* and P450<sub>Bm-3</sub> proteins are classified as fatty acid hydroxylases (Schunck et al., 1991; Seghezzi et al., 1992; English et al., 1996; Lottermoser et al., 1996). *CYP52E1* and *E2* nucleic acid sequence identities ranging between 82 and 84%, indicate that the genes are non-allelic and even though both exhibit high primary structure homologies, both constitute a suitable model for the investigation of structure-function relationships (Lottermoser et al., 1996). Should the *C. apicola* P450s hydroxylate long-chain fatty acids to [ω-1]- or ω-hydroxy fatty acids, such a subterminal fatty acid hydroxylation would represent a new catalytic activity for yeast P450 proteins.

Unlike the glucose *CYP52* repression systems identified in *C. tropicalis* and *C. maltosa* (Seghezzi et al., 1992; Ohkuma et al., 1998), six of the eight *CYP52* genes isolated from *Y. lipolytica* appear to be repressed by glycerol (Iida et al., 1998 & 2000). The first of the *Y. lipolytica* *CYP52* genes, designated *CYP52FI(YIALK1)*, was observed to be highly induced by decane, dodecane and tetradecane (Iida et al., 1998). No signal was observed for myristic acid and it was therefore suggested that the main function of *YIALK1* had no part in the omega-hydroxylation of fatty acids. A *YIALK1* disruptant showed poor growth on decane, but grew well on the longer-chain n-alkanes such as hexadecane. Analysis of the remaining seven genes, *YIALK2*-YIALK8, showed five of the seven to be induced by n-alkanes. Disruption of *YIALK2, YIALK3, YIALK4* and *YIALK6* showed almost no change in growth on media containing n-alkanes of various chain lengths (Iida et al., 2000). It was concluded from further disruption analysis that both *YIALK1* and *YIALK2* functioned primarily to assimilate n-alkanes longer than decane since a strain with both
YlALK1 and YlALK2 disrupted, was unable to grow on hexadecane. The remaining YlALKs appeared to have little significance in the assimilation of C_{10}-C_{16} n-alkanes (Iida et al., 2000). It was concluded that YlALK3-YlALK6 might be involved in the assimilation of molecules longer than hexadecane, whilst YlALK7 and YlALK8 were not induced under any of the conditions tested (Iida et al., 2000).

Finally, the appearance of peroxisomes plays an important role in alkane and/or fatty acid assimilation. Peroxisomes of the yeast C. tropicalis have been reported to be induced by oleic acid or n-alkanes (C_{10}-13), and are active in the β-oxidation of long-chain fatty acids (Kamiryo & Okazaki, 1984). In yeast, cellular fatty acid degradation into acetyl-CoA units occurs exclusively in peroxisomes (Fukui & Tanaka, 1981). The enzymes that contribute to β-oxidation in yeast include, in order, acyl-CoA oxidase, multifunctional enzyme type 2 (MFE2) [with 2-enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase activities], and the cleavage enzyme, 3-ketoacyl-CoA thiolase (Qin et al., 2000).

Enzymes involved in peroxisomal β-oxidation are synthesized on polysomes free in the cytosol, and are targeted to peroxisomes by cis-acting peroxisomal targeting signals (PTSS) [Marshall et al., 1995]. Such peroxisome targeting signals include PTS1, PTS2, mPTS, as well as a subset of trans-acting proteins involved in peroxisomal protein transport called peroxines (Smith et al., 2000). Peroxins are required for protein assembly, while subsets of peroxins are required for the targeting and import of peroxisomal matrix proteins. In yeast, the size and number of peroxisomes are regulated by Pex1p, Pex11p, Pex16p, and Pmp47 in response to growth on carbon sources metabolized by peroxisomes (Nakagawa et al., 2000; Smith & Rachubinski, 2000). Here Pmp47, a peroxisomal membrane protein, was found to be essential in middle-chain length fatty acid metabolism as well as peroxisomal proliferation (Nakagawa et al., 2000). The regulation of peroxisomal size and number in Y. lipolytica has also been attributed to the β-oxidation of fatty acids (Smith et al., 2000). Of importance here is the peroxisomal β-oxidation multifunctional enzyme type 2 (MFE2). The inability of MFE2 to function in β-oxidation leads to enlarged peroxisomes, thereby affecting the peroxisome number. Another point to note is that the Y. lipolytica MFE2 gene contains a putative upstream oleic acid-response element, as do the genes of C. tropicalis and S. cerevisiae (Einerhand et al., 1993; Smith et al., 2000). Work done by Qin et al. (2000), suggests a common ancestor for the yeast MFE2 and mammalian peroxisomal multifunctional enzyme type 2 (perMFE2) genes. Despite the fact that the yeast peroxisomal and mammalian mitochondrial and peroxisomal β-oxidation enzymes are different, similarities are evident when comparing the effects specific substrates have on peroxisomal proliferation. Fatty acids are known to promote peroxisomal proliferation, but they have also been reported to induce monoxygenases in the CYP4A gene subfamily, the B. megaterium cytochrome P450_{B3A1,3}, as well as the CYP52 genes present in C. tropicalis, C. maltosa, C. apicola and Y. lipolytica (Ramana & Kohli, 1998, Seghezzi et al., 1992, Lottermoser et al., 1996; Ohkuma et al., 1998).

In higher eukaryotes, specific DNA sequence motifs located upstream of PP target genes (PPRE), are responsible for the activation of β-oxidation enzymes and CYP4A. A similar
mechanism exists for the induction of cytochrome P450BM-3 in *B. megaterium* (English *et al.*, 1996). Here, fatty acids can interact with a repressor molecule, inducing cytochrome P450BM-3 gene expression. In both instances, upstream regulating elements are actively involved in the transcription of different genes in response to fatty acids. What this means is that there must be common upstream regulating elements occurring within the yeast and mammalian β-oxidation genes as well as the CYP52 genes, due to the fact that they are induced by similar substrates. Elucidating the fungal and mammalian transcriptional activating sequences present in the β-oxidation and CYP52 enzyme systems could lead to an understanding of the regulation of enzymes that contribute to β- and terminal oxidation reactions occurring within eukaryotic organisms.

2.6. Conclusions

The cytochrome P450 superfamily comprises more than 150 P450 genes present in a variety of organisms (Nelson *et al.*, 1993). The genes have been proposed to have evolved during "animal-plant warfare" as a protection mechanism by which animals battled against both dietary and environmental toxins (Gonzalez & Nebert, 1990). From an evolutionary standpoint, the induction of P450s resulting in enhanced xenobiotic detoxification, would most likely be advantageous. On the other hand, whether regulatory pathways evolved from regulatory pathways already present, or whether they evolved to deal with environmental and dietary xenobiotics, remains unclear. For example, should endogenous ligands regulate cytochrome P450 transcription, then from an evolutionary standpoint, this could be interpreted as a mechanism by which the need for cytochrome P450 gene regulation was to maintain metabolic homeostasis.

Initial studies focusing on endogenous metabolism by mammalian P450 enzyme systems, concentrated on pathways producing steroid hormones from cholesterol in the adrenal cortex and gonads. Consequently, P450s necessary for vitamin D metabolism in the kidney and bile acid biosynthesis in the liver, were identified. Contrary to xenobiotic metabolism, P450s involved in such endogenous metabolic systems are located in separate subcellular compartments. The microsomal P450s, as well as those localized to the inner mitochondrial membrane, are involved in bile acid biosynthesis, steroid hormone biosynthesis and vitamin D metabolism. According to the P450 nomenclature system (Nelson *et al.*, 1993), those P450s involved in the above-mentioned biosyntheses, constitute members distinct from one another as well as from those involved in xenobiotic metabolism. This being the case, the existence of P450 gene families containing forms known to be involved in xenobiotic metabolism, have also been found to contain members required for the production of key endogenous compounds.

The xenobiotic-metabolizing mammalian cytochrome P450s are found within the CYP1, CYP2, CYP3 and CYP4 gene families. The regulation of these enzymes occurs predominantly at the transcriptional and post-transcriptional level (Ramana & Kohli, 1998; Honkakoski & Negishi, 2000). The expression of the individual P450 gene families is enhanced either by receptor, or nonreceptor-mediated mechanisms. Studies have shown that PAH, glucocorticoids and PP regulate the CYP1A, CYP3A and CYP4A genes by way of receptor-mediated mechanisms. Both ethanol and PB have been shown to follow

Cytochrome P450_{B3} is a fatty acid-inducible P450 found in the prokaryote *B. megaterium* (Narhi & Fulco, 1987). It is the most catalytically efficient cytochrome P450-dependent monooxygenase known, and exhibits highest sequence identity with members of the mammalian CYP4A gene subfamily (Denison & Whitlock, 1995). Cytochrome P450_{B3} is induced by a range of structurally diverse peroxisome proliferators, and its resulting response can be compared to the potency of the drugs in mammals (English et al., 1996). Similarly, a second prokaryotic cytochrome P450 monooxygenase extensively studied in *P. putida*, constitutes P450_{cam}. The *P. putida* CAM plasmid encodes oxidative enzymes for a pathway for the catabolism of camphor to isobutylate (Koga et al., 1986; Fujita et al., 1993; Aramaki et al., 1993), and has long been the prototype for the prokaryotic P450s. Extensive physical and biochemical characterization of the enzyme, paved the way for the use of rational site-directed mutagenesis to further probe the structure-function relationships in P450_{cam}.

Cytochrome P450 multiplicity, as is seen for mammalian systems, is rarely found in microorganisms. Yeasts harboring cytochrome P450 P450 multigene families include *C. maltosa*, *C. tropicalis*, *Y. lipolytica*, and to a lesser extent, *C. apicola*. The n-alkane assimilating yeasts *C. maltosa* and *C. tropicalis* contain multiple n-alkane-inducible forms of P450 (P450_{alk}). Eight structurally related P450_{alk} genes belonging to the CYP52 family have been identified for *C. maltosa*, while *C. tropicalis* has been shown to contain seven CYP52 genes (Schunck et al., 1989 & 1991; Ohkuma et al., 1991 & 1998; Seghezzi et al., 1991 & 1992). The CYP52 multigene families found in *C. maltosa* and *C. tropicalis* both exhibit glucose repression systems. In general, the genes are not only induced by hydrocarbon substrates, but the P450 enzymes secreted are actively involved in the hydroxylation of n-alkanes and fatty acids alike (Seghezzi et al., 1992; Zimmer et al., 1998; Ohkuma et al., 1998). Much the same can be said for the CYP52 multigene family found in *Y. lipolytica* (Iida et al., 1998 & 2000). Six of the eight genes present are induced by various n-alkanes, but additional studies are required to determine whether the proteins produced are active. Unlike the *C. maltosa* and *C. tropicalis* CYP52 genes, the *Y. lipolytica* CYP52 genes are not repressed by glucose, but by glycerol (Iida et al., 2000). The two CYP52 genes isolated from the sophorose lipid-producing yeast *C. apicola* (Lottermoser et al., 1996) appear to be induced by long-chain fatty acids. The proteins produced show high homology with the fatty acid hydroxylases, namely the mammalian CYP4A and *B. megaterium* P450_{B3} proteins. Additional fungal P450s found to share characteristic properties with other phyla include the CYP51 gene product, sterol 14a-demethylase (P45014 DM) [Venkateswarlu et al., 1995; Lamb et al., 1995]. The CYP51 family is the oldest eukaryotic CYP family known at present, indicating the essential role of sterol biosynthesis in eukaryotes.

In recent years the application of the techniques of genetic engineering and molecular biology to the study of mixed-function cytochrome P450s, has led to new insights into the mechanism of their regulation. When considering the large number of isoenzymes in the CYP gene superfamily, it is not surprising that the different genes employ different...
regulatory strategies when comparing their evolutionary history and wide range of functions. Advances in the field of molecular biology will help to clarify the regulatory complexities exhibited by this superfamily of genes, allowing us to further understand the interactions occurring between the various regulatory mechanisms abundant in an entire organism.
Chapter Three

Materials and Methods

3.1 Materials

3.1.1 Strains, media and growth conditions

A wild-type strain *Yarrowia lipolytica* UOFS Y-0164 obtained from Dr. J.-M. Nicaud as W29 was cultured on yeast extract/peptone/glucose (YEPD) [IL contained: 20g glucose (Saarchem), 20g peptone (Biolab), 10g yeast extract (Biolab)]. One x yeast nitrogen base (YNB) consisted of 0.67g YNB dissolved in 100ml distilled water, while 2xYNB consisted of 0.67g YNB dissolved in 50ml distilled water.

(i) Induction experiments using glucose pre-grown cells: A 100ml *Y. lipolytica* UOFS Y-0164 SD (2% glucose; 1xYNB) pre-inoculum in 1L shake flasks was grown at 30°C for 16hrs on a rotary shaker (600rpm) from a 24hr YEPD plate. Cells were spun down at 1000xg for 5min, washed in 1xYNB medium, and pelleted at 1000xg for 5min. Cells were then resuspended in 2xYNB from which 5ml aliquots were transferred to the 10ml inducing medium. For induction of cytochrome P450s, the yeasts were grown at 30°C in YNB (Difco) supplemented with 2% *n*-alkane/fatty acid (Aldrich). Hydrocarbons used for cytochrome P450 induction were polyoxyethylene sorbitan monooleate (Tween 80), 2,6,10,14-tetramethylpentadecane (pristane), glucose, tetradecane (C14), hexadecane (C16), octadecane (C18), docosane (C22), octacosane (C28), stearic acid (C18OOH), and behenic acid (C22OOH). (ii) Induction experiments using acid pre-grown cells: A 10ml pre-culture in a medium containing 2% peptone, 1% yeast extract, 2% acid (C18OOH/ C22OOH) was grown at 30°C for 24 hours. Cells were spun down at 1000xg for 5min, washed in an equal volume of cyclohexane, and spun down again at 1000xg and any further occurrence for 5min. Cells were then resuspended in 100ml culture media containing 0.5M phosphate buffer (pH 6.8), 0.1% yeast extract, 1xYNB with amino acids, and 2% acid (C18OOH/ C22OOH). Cells were grown at 30°C for 16 hours, spun down at 1000xg for 5min, washed in an equal volume of 1xYNB, spun down again at 1000xg for 5min. Cells were then resuspended in 2xYNB from which 5ml aliquots were transferred to the 10ml inducing medium constituting both 2% C18OOH and 2% C22OOH. Fatty acids were emulsified by sonicating 5g of acid and 0.5ml Tween 80, in distilled water (final volume 50ml) for 3 min.

*E. coli* JMB109 (*recA1, endA1, gyrA96, thi, hsdR17* (rK-, mK+), *relA1, supE44, Δ(lac-proAB), [F, *traD36, proAB, lacI^qZAM15)]) (Promega), used for plasmid preparations, was grown in SOC [IL contained: 10ml 2M glucose (Saarchem), 5g yeast extract (Biolab), 20g tryptone (Biolab), 1ml 2M MgSO4 stock, 2.5ml 1M KCl, 10ml 1M NaCl (Saarchem)] or LB media as defined by the suppliers of the kit.
3.1.2 Chemicals, enzymes and kits

All chemicals were of analytical or molecular biology grade and were used without further purification. Sorbitol, sodium citrate, sodium dodecyl sulfate (SDS), ethylenediaminetetraacetic acid (EDTA), sodium acetate and guanidine hydrochloride were obtained from Merck, while YNB (with amino acids and ammonium sulfate), was obtained from Difco. Two-amino-2-(hydroxymethyl)-1,3-propanediol (Tris) and 4-morpholinepropanesulfonic acid (MOPS) were obtained from Roche, while Tween 80 and 2-mercaptoethanol were obtained from Sigma. Hexadecane, octacosane, stearic acid and 2-nitrophenyl-β-D-galactopyranoside were obtained from Aldrich. Tetradecane, octadecane, docosane, behenic acid and pristane were obtained from Fluka. Proteinase K, bovine serum albumin (BSA), RNase, Taq DNA polymerase, Lyticase, Ampicillin, EcoRI, β-agarase and High Prime labeling kit were obtained from Roche. pGEM-T® Easy Vector, Wizard® PCR Preps DNA Purification kit, RQ1 RNase-Free Dnase and Access RT-PCR® Introductory System were obtained from Promega. Thermo Sequenase™ version 2.0 DNA sequencing kit and [α-32P]dCTP were obtained from Amersham while GeneScreen Plus™ membranes used were obtained from Nen® Life Science. Primers used were synthesized by Whitehead Scientific and Roche.

3.2 Methods

3.2.1 Isolation of chromosomal DNA

The yeast cells were retrieved from 50ml YEPD liquid cultures grown in 250ml shake flasks at 30°C to an O.D600 of 1.0. Yeast cells were harvested by centrifugation at 1000xg for 5min using a Beckman GS-15R centrifuge. Each cell pellet was resuspended in 10ml 1M Sorbitol and repelleted. The cells were resuspended in 150μl 1M Sorbitol/100mM sodium citrate (pH 7.8) and transferred to 1.5ml eppendorf tubes. Ten microlitres of 5mg/ml Lyticase was then added and incubated at 30°C for 2 hours. Light microscopy was used for the identification of spheroplasts. The spheroplasts were centrifuged at 10000xg and elsewhere for 5 seconds and the supernatant was discarded as thoroughly as possible. One hundred and fifty microlitres of 4.5M GuHCl, 100mM EDTA, 150mM NaCl, 0.05% (w/v) Sarkosyl (pH 8.0) was added and the spheroplasts were mixed gently using a pipette tip. The spheroplast sample was incubated at 65°C for 10 min, after which it was cooled to 25°C. One hundred and fifty microlitres of absolute ethanol (-20°C) was added and the mixture was microcentrifuged at 10000xg for 5min at 4°C. The pellet well was drained and 300μl 10 x TE was added. The pellet was then disrupted and vortexed gently until the pellet was resuspended. Ten microlitres of 10mg/ml Proteinase K was added and incubated at 65°C for 30min. An equal volume of 1:1 chloroform/phenol mixture was added, the mixture was vortexed and centrifuged at 10000xg for 2min. This was repeated once. An equal volume of chloroform/iso-amylalcohol was added, the mixture was then vortexed and centrifuged at 10000xg for 2min. The aqueous layer was transferred to a new eppendorf tube to which 10% of 3M sodium acetate (v/v) was added. The DNA was precipitated with two volumes of absolute ethanol at -70°C for 15min. The precipitated DNA was pelleted by centrifugation at 10000xg for 10min. The DNA was washed in
70% ethanol and resuspended in 50µl 1 x TE. RNA was digested with 1µl of 10mg/ml RNase (Roche) at 37°C for 60min.

3.2.2 Isolation of total RNA

The yeast cells were induced in liquid medium with either n-alkane or fatty acid as carbon source, and were harvested as described in section 3.1.1. Induction times on the carbon source of interest ranged from five to thirteen hours and all yeast cells were induced at 30°C. Ten millilitres of the culture (5-8 x 10⁷ cells/ml) were centrifuged as previously described (section 3.1.1), resuspended in 1.5ml of YNB without carbon source, and transferred to a 2.2ml microfuge tube. The following two methods were used for RNA isolation: The first method was done according to the procedure described by De Winde et al. (1996). Yeast cells were resuspended in 0.5ml extraction buffer (111111111111 EDTA; 0.1M LiCl; 0.1M Tris; 10mM sodium acetate, pH 7.5). Sterile glass beads, 0.6ml phenol/chloroform/iso-amylalcohol (PCI) [equilibrated with H₂O] and 100µl of 10% SDS were then added to the cells. The mixture was vortexed vigorously until the pellet was resuspended and then chilled on ice. The solution was centrifuged at 1000xg for 5min at 4°C after which the aqueous phase was removed. A second aliquot of PCI was then added and the extraction step was repeated until no interphase was visible. The aqueous layer was transferred to a new eppendorf tube and the salt concentration was adjusted to 0.4M using sodium acetate. The RNA was precipitated with two volumes of absolute ethanol at -20°C for 20 hours. The precipitated RNA was then pelleted by centrifugation at 10000xg for 15min at 4°C. The RNA was washed in 70% ethanol and resuspended in 15µl Nuclease-free water. Contaminating DNA was digested using RQ1 Rnase-Free Dnase according to the manufacturer’s recommendations. The second method was followed according to the procedure described by Domdey et al. (1984). Following centrifugation at 1000xg for 5min, 0.7ml of AE buffer (50mM sodium acetate, pH 5.3; 10mM EDTA), 100µl of 10% SDS and 1ml of fresh phenol (equilibrated with AE buffer), was added to the cells. The pellet was disrupted and vortexed vigorously until the pellet was resuspended. The mixture was then incubated at 65°C for 4min and rapidly chilled in a dry ice-ethanol bath until phenol crystals appeared. The frozen mixture was centrifuged at 10000xg for 10min at 4°C after which the phenol phase was removed leaving the interphase intact. A second aliquot of phenol was then added and the extraction step was repeated once. The aqueous layer was transferred to a new eppendorf tube. An equal volume of 1:1 phenol/chloroform mixture was added, the mixture was vortexed and centrifuged at 10000xg for 15min at 4°C. The aqueous layer was transferred to a new eppendorf tube and the salt concentration was adjusted to 0.4 M LiCl (pH 5.3). The RNA was precipitated with two volumes of absolute ethanol at -20°C for 20 hours. The precipitated RNA was then pelleted by centrifugation at 10000xg for 15min at 4°C. The RNA was washed in 70% ethanol and resuspended in 15µl Nuclease-free water. Contaminating DNA was digested using RQ1 Rnase-Free Dnase according to the manufacturer’s recommendations.
3.2.3 Primers for amplifying CYP52 and Actin gene fragments from \textit{Y. lipolytica} UOFS Y-0164

The specific \textit{Y. lipolytica} P450\textit{alk} primers were based on the non-conserved 5'-regions of the eight CYP52 genes present in the yeast. The CLUSTALW program was used for multiple alignment of the nucleotide sequences of the genes. The default settings have: pair gap of 0.05; end gap of 10; gap extension of 0.05; gap distance of 0.05 and gap open of 10. Eight different upstream primers (ALK-1 to ALK-8) represented each gene individually while a common downstream primer, (ALK-REV) was also designed:

ALK1: 5' - GCCCTCAACCTGCGCCTCGCCTCG-3'  
ALK2: 5' - TTTATACCTTTTGCTGCTGCTGCT-3'  
ALK3: 5' - ATCGAAACGCTCATTGAGCGGTCG-3'  
ALK4: 5' - TGGCCGCTCTGTTGTGGCCCTGCA-3'  
ALK5: 5' - AATGATCCAGTCGGTTTTCTTGGCCT-3'  
ALK6: 5' - ATCGATCTGCTGCCCTTTGCC-3'  
ALK7: 5' - CTAGGTACCTGCTGCCCTTTGCA-3'  
ALK8: 5' - TACTACTACTACCCCTACAGACAG-3'  
ALK REV: 5' - G[AGT]G[GT][GC][GC][CT][AG][AG][GCT][GC][CT]TTTGA[GC][CT]-3'

For quantitative purposes a set of actin primers were designed. Open reading frames coding for the actin gene from \textit{Rattus norvegicus} (rat), \textit{Gallus gallus} (chicken), \textit{Homo sapiens}, \textit{Candida apicola} and \textit{Saccharomyces cerevisiae} were aligned using CLUSTALW. The default settings are the same as above. The oligonucleotides are constitute an actin upstream primer [ACT(F)] and a downstream primer [ACT(R)]:

ACT (F): 5' - CCAACTGGGACGATATGGTTCTGGC-3'  
ACT (R): 5' - GAAGCCAAGATAGAAACCAATCCAGACAGTA-3'

3.2.4 PCR amplification of CYP52 gene fragments from \textit{Y. lipolytica} UOFS Y-0164 using ALK-1/ALK-REV to ALK-8/ALK-REV primer pairs

The PCR amplification of the isolated \textit{Y. lipolytica} genomic DNA was performed under standard reaction conditions. The PCR reaction mixture of 50\textmu l contained 1\textmu l of 0.668\textmu g/\textmu l template DNA, 5\textmu l of each 5\textmu M primer, 5\textmu l of 2.0mM deoxyribonucleoside triphosphate and 5\textmu l of 10 x polymerase buffer containing 15mM MgCl\textsubscript{2}, 0.5\textmu l of 5 U/\textmu l Taq DNA polymerase (Roche). The PCR programme consisted of an initial DNA denaturation at 94°C for 2min followed by 25 cycles of denaturation at 94°C for 1min, annealing at 55°C for 1min and primer extension at 72°C for 1min. Primer extension was then continued at 72°C for 7min.

3.2.5 PCR amplification of Actin gene fragments from \textit{Y. lipolytica} UOFS Y-0164 using ACT(F)/ACT(R) primer pair

The PCR reaction mixture of 50\textmu l contained 1\textmu l of 0.338\textmu g/\textmu l template DNA, 1\textmu l of each 20\textmu M primer, 4\textmu l of 2.5mM deoxyribonucleoside triphosphate, 5\textmu l of 10 x polymerase buffer containing 15mM MgCl\textsubscript{2} and 1\textmu l of 5 U/\textmu l Taq DNA polymerase (Roche). The PCR programme consisted of an initial DNA denaturation at 94°C for 2min.
followed by 25 cycles of denaturation at 94°C for 1min, annealing at 55°C for 1min and primer extension at 72°C for 1min. Primer extension was then continued at 72°C for 7min.

3.2.6 Cloning and DNA sequencing

PCR bands were extracted from 1% low-melting agarose gels using β-Agarase (Roche) according to the manufacturer’s recommendations. The purified PCR products were ligated into the pGEM®-T Easy Vector system (Promega) according to the manufacturer’s recommendations. Transformation of the E. coli JMB109 competent cells was done by electroporation at 2.5kV, 25μF and 200Ω using the Gene Pulser (Biorad). Clones were selected by blue-white screening on LB agar plates containing 100μg/ml ampicillin, 10μg/ml isopropyl-1-thio-β-D-galactoside (IPTG) and 40μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). Plasmid DNA was isolated from positive clones by alkaline lysis miniprep (Sambrook et al., 1989). The sizes of the cloned fragments were confirmed by EcoRI (Roche) restriction analysis. Inserts in the pGEM®-T Easy Vector system were sequenced directly with Y. lipolytica eYP52 primer pairs. The sequencing was based on the dideoxynucleotide chain termination method using fluorescence-labeled dye terminators. Inserts were labeled using the Sequenase™ Version 2.0 DNA Sequencing Kit (Amersham) according to the manufacturer’s recommendations. The elongation products from the sequencing reactions were analysed using ABI PRISM™ 377 automatic sequencer and Sequence Navigator (Perkin Elmer).

3.2.7 RT-PCR amplification of mRNA from Y. lipolytica UOFS Y-0164 grown on long-chain n-alkanes/fatty acids using ALK-1/ALK-REV to ALK-8/ALK-REV primer pairs

The RT-PCR reactions were performed using the Access RT-PCR System (Promega). The RT-PCR reaction mixture of 50μl contained 10μl of AMV/T3 5x Reaction Buffer, 1μl of 10 mM deoxynucleoside triphosphate mixture, 1μl of each 20μM primer pair (ALK-1/ALK-REV to ALK-8/ALK-REV), 2μl of 25mM MgSO₄, 0.5μl of AMV Reverse Transcriptase (5U/μl), 0.5μl of T3 DNA Polymerase (5U/μl), 1μl template RNA and nuclease-free water. The reverse transcription was performed at 48°C for 45min. The PCR programme consisted of an initial DNA denaturation at 94°C for 2min followed by 35 cycles of denaturation at 94°C for 30s, annealing at 55°C for 1min and primer extension at 68°C for 1min. Final primer extension was at 68°C for 7min. Negative control reactions for each RT-PCR reaction constituted replica RT-PCR reaction mixtures without the AMV Reverse Transcriptase. Five microlitre samples were run on 1.5% (w/v) agarose gels containing 0.5μg/ml ethidium bromide.

3.2.8 Northern hybridizations

Total RNA of Y. lipolytica cells was extracted as described in section 3.2.2 (Method 2) and electrophoresed in 1% agarose gel containing 1x MOPS (20mM MOPS, pH 7.0; 5mM Sodium Acetate; 1mM EDTA) and 6% formaldehyde. Ethidium bromide was
added at a concentration of 0.5µg/ml to allow direct UV-visualization of ribosomal RNA bands. Total RNA was transferred to Gene Screen Plus™ membrane by electoblotting (Hoefer Scientific Instruments) at 700mA for 24hrs as specified by the manufacturer’s instructions. The RNA was crosslinked to the membrane using a GS Gene Linker™ UV Chamber (Bio-Rad). Clones containing \textit{Y. lipolytica CYP52} gene fragments were digested with \textit{EcoRI} and concurrently run on 1% (w/v) low-melting agarose gels containing 0.5µg/ml ethidium bromide. Probes were purified using the Wizard® PCR Preps DNA Purification System according to the manufacturer’s recommendations. DNA probes (80.6ng/µl) were labeled with [\alpha-^{32}P]dCTP (Amersham) using High Prime labeling kit (Roche) according to the manufacturer’s recommendations. Unincorporated dNTPs were removed using the Wizard® DNA PCR Preps Purification System (Promega) according to the manufacturer’s recommendations. Membranes were prehybridized (1%BSA; 0.525M Na₂PO₄, pH 7.2; 7% SDS; 1mM EDTA) for 1 hour at 68°C and then hybridized (1%BSA; 0.525M Na₂PO₄, pH 7.2; 1%SDS; 1mM EDTA) for 24hrs at 68°C. The hybridization membranes were washed twice at R/T in low stringency solution (0.04M Na₂PO₄, 1mM EDTA; 1% SDS) for 10min, followed by two wash cycles at 65°C for 30min. A high stringency washing solution (0.1x SSC; 0.1% SDS) at 65°C for 30min was used where necessary. Membranes were exposed over CP-G Plus medical X-ray film (AGFA) at -80°C for 20hrs. Probes were removed by pouring a boiling 0.1% SDS solution over the membranes. A comparison of the signals generated by the different probes on different substrates was achieved using an Actin internal standard.

\subsection*{3.2.9 Southern hybridizations}

For Southern hybridizations 80.6ng/µl of purified probes were separated on a 1% TAE agarose gel and then transferred to Gene Screen Plus™ membrane by electoblotting (Hoefer Scientific Instruments) at 700mA for 24hrs as specified by the manufacturer’s instructions. The DNA was crosslinked to the membrane using a GS Gene Linker™ UV Chamber (Bio-Rad) Transfer, hybridization, exposure and washing were done as described for Northern hybridizations.

\subsection*{3.2.10 Sequence analysis}

The automated nucleotide and amino acid sequence alignments were done using CLUSTALW (http://www2.ebi.ac.uk/clustalw) and DNAAssist Version 1.02 (Patterson & Graves, 2000). Phylogenetic analyses of all eight \textit{Y. lipolytica} and \textit{Candida CYP52} forms were created using Phylip Parsimony Algorithm Version 3.572c.
Chapter Four

Results and Discussion

4.1 PCR amplification of CYP52 and Actin gene fragments from Y. lipolytica UOFS Y-0164

The eight available Y. lipolytica CYP52 gene sequences were aligned using CLUSTALW. Eight gene-specific primers (ALK-1 to ALK-8) were designed according to the nucleotide sequences of the variable 5' -region of the CYP52 genes, while a common downstream primer (ALK-REV) was designed according to homologous regions exhibited between the 8 CYP52 genes (Fig. 4.1A). For quantitative purposes actin was used as an internal control. At the time, the Y. lipolytica Actin gene sequence was not available. Subsequently, sequences corresponding to the open reading frames for the Actin genes from Gallus gallus (chicken), Rattus norvegicus (rat), Candida apicola, Saccharomyces cerevisiae and Homo sapiens were selected and aligned using CLUSTALW (Fig. 4.1B). The corresponding alignment allowed the design of the Actin primer pair ACT(F)/ACT(R). Amplification of the genomic DNA of Y. lipolytica UOFS Y-0164 using the CYP52 primer pairs ALK-1/ALK-REV to ALK-8/ALK-REV and the actin primers ACT(F) and ACT(R) generated the expected products of approximately 329 and 851bp in length, respectively (Fig. 4.1C).

Figure 4.1C. PCR amplification of Y. lipolytica UOFS Y-0164 CYP52 and Actin gene fragments. Molecular size marker (lane 1): EcoRI/HindIII-digested λ phage ladder. Products of PCR reactions using primer pairs ALK-1/ALK-REV (lane 2); ALK-2/ALK-REV (lane 3); ALK-3/ALK-REV (lane 4); ALK-4/ALK-REV (lane 5); ALK-5/ALK-REV (lane 6); ALK-6/ALK-REV (lane 7); ALK-7/ALK-REV (lane 8); ALK-8/ALK-REV (lane 9); and ACT(F)/ACT(R) [lane 10].

Figure 4.1A. Alignment of the 8 CYP52 genes present in Y. lipolytica. Eight specific primers (ALK-1 to ALK-8) were designed from the 5'-region of the CYP52 genes. The design of the common downstream primer (ALK-REV) was based on homologous nucleic acid sequences exhibited between all 8 CYP52 genes. Homology is indicated by asterisks and the primers by bold, italicized and underlined font. Alignment performed using CLUSTALW.
**ALK(1-8) primers**

### ALK-JS primers

- **ALK-JS F:** ATGATCCAGTCGGTTTTCTTGGCCTTGGC
- **ALK-JS R:** ATGTTGACCAATCTTACGATCGTGCTGACCC

### ALK-GS primers

- **ALK-GS F:** ATGTTCCAGCTATTCTC-GATACTGGT-GCTGGCTTTCACA
- **ALK-GS R:** ATGCTACAACTCTTTGG-CGTCCTTGT-GTTGGCGCTGA

### ALK-REV primer

- **ALK-REV F:** ATGATCCAGCTTCCTCATTACACTCTACTACTACTACTACCC-TTACTGATCATCCC
- **ALK-REV R:** CATTACGCTCTTTTTGCTAGTACTAAGCGTTCAAGTCGACACTGTTCATAGACGTTACA-

### ALK-REV primer

- **ALK-REV F:** GTACAGACATGTCTAACAAGGCTACATACGCTGCAAGCTGACCATAC-TAC---CTCTACATCA
- **ALK-REV R:** GTAACAGATGCTACA.CVGGAGTGGAGAGCACTGGAGACCT-GCTGCTGCTGCACTC

**ALK-REV' primer**

- **ALK-REV F:** GTAAGGCTGTGGCATGCCTCCCATGGCAA.TGGAGGGTTT-CTGGGTTAGGCCA
- **ALK-REV R:** TGCCGCTCTTCTTCCTCATCATCA.GCTGCTCGAACACCT.

---

**52**
Figure 4.1B. Alignment of the Actin open reading frames (ORF) present in Rattus norvegicus, Homo sapiens, Gallus gallus, Candida apicola and Saccharomyces cerevisiae. Two conserved primers were designed: An upstream primer [ACT(F)] based on the 5'-region of the Actin ORFs as well as a downstream primer [ACT(R)]. Homology is indicated by asterisks and the primers by bold, italicized and underlined font. Alignment performed using CLUSTALW.

4.2 Cloning and sequencing of CYP52 PCR gene fragments from Y. lipolytica UOFS Y-0164

The individual PCR products were ligated into pGEM®-T Easy Vector and the ligation mixture was transformed into E. coli cells. Ten positive colonies were grown in LB-amp media and the plasmids were isolated using the alkaline lysis method (Sambrook et al., 1989). Digestion of the plasmids with EcoRI confirmed the presence of the insert in the vector (Fig. 4.2).

Figure 4.2. Cloning of CYP52 Y. lipolytica UOFS Y-0164 PCR gene fragments. 1.5% (w/v) agarose gel containing 329bp EcoRl-digested pGEM®-T Easy Vector inserts. 1kb ladder (lanes 1 & 22). 329bp YIALK1 PCR gene fragments (lanes 2 & 3); YIALK2 (lanes 4 & 5); YIALK3 (lanes 6 & 7); YIALK4 (lanes 8 & 9); YIALK5 (lanes 10, 11 & 12); YIALK6 (lanes 13, 14 & 15); YIALK7 (lanes 16, 17 & 18) and YIALK8 (lanes 19, 20 & 21).
The *Y. lipolytica* CYP52 specific primer pairs ALK-1/ALK-REV to ALK-8/ALK-REV were used to sequence pGEM®-T Easy Vector inserts from selected plasmids. Comparison of the *Y. lipolytica* CYP52 gene sequences (Accession Numbers AB010388-AB010395) with the positive clones obtained confirmed the identity of the individual CYP52 sequences (Table 4.1).

**TABLE 4.1. COMPARISON OF THE NUCLEIC ACID Y. lipolytica CYP52 GENE AND CLONAL* SEQUENCES**

<table>
<thead>
<tr>
<th>Identity/Similarity</th>
<th>Clone1</th>
<th>Clone2</th>
<th>Clone3</th>
<th>Clone4</th>
<th>Clone5</th>
<th>Clone6</th>
<th>Clone7</th>
<th>Clone8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alk1</td>
<td>72</td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>8</td>
<td>10</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Alk2</td>
<td>12</td>
<td>90</td>
<td>7</td>
<td>11</td>
<td>11</td>
<td>7</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Alk3</td>
<td>7</td>
<td>8</td>
<td>80</td>
<td>17</td>
<td>37</td>
<td>7</td>
<td>35</td>
<td>7</td>
</tr>
<tr>
<td>Alk4</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>90</td>
<td>32</td>
<td>52</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Alk5</td>
<td>5</td>
<td>13</td>
<td>14</td>
<td>32</td>
<td>91</td>
<td>37</td>
<td>60</td>
<td>5</td>
</tr>
<tr>
<td>Alk6</td>
<td>10</td>
<td>8</td>
<td>11</td>
<td>34</td>
<td>37</td>
<td>94</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Alk7</td>
<td>6</td>
<td>12</td>
<td>9</td>
<td>12</td>
<td>57</td>
<td>13</td>
<td>86</td>
<td>4</td>
</tr>
<tr>
<td>Alk8</td>
<td>5</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>61</td>
</tr>
</tbody>
</table>

*The sequence comparison was done using the CLUSTALW program at [http://www2.ebi.ac.uk/clustalw](http://www2.ebi.ac.uk/clustalw) (Matrix: blosum; Gap open: 10; End gaps: 10; Gap extension: 0.05; Gap distances: 0.05; Top diag: 1; Pair gap: 0.05). The values are given in % (amino acid similarity/identity). All clones contained the designated +/- 329bp PCR product (Fig. 4.1C) and were aligned with the corresponding *Y. lipolytica* CYP52 genes (AB010388-AB010395).*

### 4.3 Growth of *Y. lipolytica* UOFS Y-0164 on different hydrocarbons

To investigate conditions for the induction of all eight *Y. lipolytica* CYP52 genes, growth was monitored to determine when total RNA should be isolated (Fig. 4.3). Due to the insolubility of the hydrophobic long-chain alkanes (>C20) and long-chain acids, the use of a co-solvent or surfactant is generally required. Growth on media containing very long-chain alkanes (>C20) or fatty acids without co-solvents or surfactants was very poor. The addition of pristane as a co-solvent significantly improved growth on the longer-chain alkanes (>C20), while solubilization of the fatty acids was achieved using Tween 80. Both pristane and Tween 80, when used as a sole carbon source (Fig. 4.3A & C), failed to produce an increase in cellular biomass. Growth of the UOFS Y-0164 strain on n-alkanes under the conditions used in these experiments was very poor. Better growth can be obtained if the cultures are pre-grown on yeast extract peptone media supplemented with the hydrocarbon of interest, and if 0.1% yeast extract is added to the YNB media. However, for comparative purposes, growth conditions employed were the same as those used by Iida *et al.* (1998).
Figure 4.3. Utilization of n-alkanes/fatty acids by Y. lipolytica UOFS Y-0164. A: (Φ) Control; (□) 2% pristane; (▲) 2% hexadecane; (■) 2% hexadecane & 2% pristane; (○) 2% octadecane; (○) 2% octadecane & 2% pristane. B: (□) 2% docosane; (▲) 2% docosane & 2% pristane; (Φ) 2% octacosane; (○) 2% octacosane & 2% pristane. C: (Φ) 0.5% Tween 80; (▲) 2% stearic acid & 0.5% Tween 80; (○) 2% benenic acid & 0.5% Tween 80.
4.4 Isolation of total RNA from *Y. lipolytica* UOFS Y-0164

The RNA isolation method described by De Winde *et al.* (1996) was initially used to isolate total cellular RNA from *Y. lipolytica* UOFS Y-0164 following induction on the *n*-alkanes/fatty acids of interest. The quality of the RNA obtained using this method was very poor for several reasons. Firstly, the glass beads used to break the cells not only released DNA from the cell nuclei, but also produced low-quality mRNA due to the shearing action of the glass beads. Secondly, the neutral pH of the extraction buffers used promoted the solubilization of DNA in the aqueous phase and resulted in the partial hydrolysis of the mRNA of interest. Initial attempts at both RT-PCR and Northern hybridization experiments yielded no results.

The RNA isolation procedure described by Domdey *et al.* (1984) produced superior quality mRNA when compared to the mRNA obtained using the De Winde method. Higher quality mRNA was obtained from the yeast cells as they were broken using hot phenol and SDS and not glass beads. The acidic pH of the extraction buffers used here favoured the solubilization of the cellular RNA in the aqueous phase and reduced the solubility of any contaminating DNA. Although this RNA isolation procedure produced superior quality RNA when compared to the method described by De Winde *et al.* (1996), mild mRNA degradation is evident as can be seen in the Northern hybridizations.

4.5 RT-PCR amplification of *Y. lipolytica* UOFS Y-0164 CYP52 gene fragments

Although simple RT-PCR can not be used for the quantification of mRNA transcripts, it can be used to determine whether specific transcripts are present or not. The eight *Y. lipolytica* CYP52 gene-specific primers were designed such that one optimized RT-PCR reaction could be used for the amplification of the CYP52 mRNA transcripts isolated from the yeast. Optimization of the initial PCR reactions (Fig. 4.1C) paved the way for further RT-PCR reactions. Total RNA was isolated from *Y. lipolytica* UOFS Y-0164 cells following five hours of induction on tetradecane, hexadecane, octadecane, pristane, and mixtures of docosane and pristane, octacosane and pristane, stearic acid and Tween 80 and behenic acid and Tween 80. Negative control reactions constituted the RT-PCR reaction performed without the reverse transcriptase (RT) step, thereby ensuring that the reaction products were not amplified from contaminating DNA. No fragments for the negative control experiments were observed for any of the substrates tested, confirming that the RT-PCR fragments obtained were derived from mRNA. 329bp ALK-1/ALK-REV gene fragments were obtained for all of the substrates tested (Fig. 4.4), indicating that the *Y*IALK1 P450 form appears to be the predominant gene involved in *n*-alkane/fatty acid assimilation. Five of the remaining seven CYP52 genes present in *Y. lipolytica* were observed to be induced on the carbon sources of interest. No mRNA transcripts were observed for *YIALK7* and *YIALK8* following five hours of induction. The resulting ALK-2/ALK-REV and ALK-3/ALK-REV primer pair PCR products appear to be less intense when compared to the intensities of the remaining CYP52 gene fragments. This could be attributed to the conditions at which the PCR reactions were optimized (Fig. 4.1C). Here, the *YIALK2* and *YIALK3* 329bp gene fragments also appear less intense when compared to the intensities of the remaining CYP52 gene fragments. As
a result, the signal intensities obtained for the RT-PCRs cannot be attributed to gene induction levels. This being the case, mRNA transcripts of six of the eight *Y. lipolytica* CYP52 genes appear to be present following induction on long-chain n-alkanes/fatty acids.

![RT-PCR amplification](image)

Figure 4.4. RT-PCR amplification of 329bp *Y. lipolytica* UOFS Y-0164 CYP52 gene fragments. Cells were induced on (i) 2% pristane; (ii) 2% tetradecane; (iii) 2% hexadecane; (iv) 2% octadecane; (v) 2% docosane & 2% pristane; (vi) 2% octacosane & 2% pristane; (vii) 2% stearic acid & 0.5% Tween 80 and (viii) 2% behenic acid & 0.5% Tween 80. 100bp Molecular size marker (lane 1); Negative control (lane 2); ALK-1/ALK-REV primer pair (lane 3); ALK-2/ALK-REV primer pair (lane 4); ALK-3/ALK-REV primer pair (lane 5); ALK-4/ALK-REV primer pair (lane 6); ALK-5/ALK-REV primer pair (lane 7); ALK-6/ALK-REV primer pair (lane 8); ALK-7/ALK-REV primer pair (lane 9); ALK-8/ALK-REV primer pair (lane 10).

4.6 The use of Northern hybridizations to monitor CYP52 gene expression in *Y. lipolytica* UOFS Y-0164

Work done initially by Iida *et al.* (1998) using the *Y. lipolytica* CX-161 strain showed that Y1ALK1 was induced by decane, dodecane, tetradecane, but repressed by glycerol. To prepare probes for Northern analysis, positive clones containing the appropriate CYP52 gene fragment were amplified using the *Y. lipolytica* CYP52 primer pairs ALK-1/ALK-REV to ALK-8/ALK-REV. The corresponding 329bp PCR products were
purified from a low-melting agarose gel and labeled as described in the Materials and Methods. To determine the specificity of the probes, 80.6 ng/µl of each 329bp CYP52 gene fragment was run on a 1% agarose gel and transferred to a positively charged nylon membrane as described in Section 3.2.9. Southern hybridization analysis showed each individual CYP52 gene probe to be specific (Fig. 4.5).

![Southern hybridization analysis](image)

Figure 4.5. Southern hybridization analysis of *Y. lipolytica* UOFS Y-0164 CYP52 gene probes using CYP52 gene-specific PCR products. 329bp ALK-1/ALK-REV PCR product (lane 1); ALK-2/ALK-REV PCR product (lane 2); ALK-3/ALK-REV PCR product (lane 3); ALK-4/ALK-REV PCR product (lane 4); ALK-5/ALK-REV PCR product (lane 5); ALK-6/ALK-REV PCR product (lane 6); ALK-7/ALK-REV PCR product (lane 7); ALK-8/ALK-REV PCR product (lane 8).

Induction of *Y*ALK1 was apparent after 5 hours following exposure to a range of substrates (Fig 4.6). Substrates tested were pristane, tetradecane, mixtures of hexadecane and pristane, octadecane and pristane, docosane and pristane, octacosane and pristane,
 Tween 80 as well as stearic acid and Tween 80. *Y. lipolytica* UOF5 Y-0164 was unable to grow on either of the co-solvents used (Fig. 4.3A & C), although both Tween 80 (Fig. 4.6) and pristane (Fig. 4.4 & 4.6) were observed to induce YlALK1. The longer-chain substrates were also observed to induce YlALK1, but due to the insolubility of the substrates in aqueous media, the inducing media contained either Tween 80 or pristane (Fig. 4.4). The induction of YlALK1 by pristane complicated matters due to the fact that *Y. lipolytica* could not grow on the longer-chain alkanes without the aid of the co-solvent. The longest alkane able to support growth of *Y. lipolytica* UOF5 Y-0164 in the absence of pristane was octadecane (Fig. 4.3A).

![Figure 4.6. Northern hybridization analysis of YlALK1 gene expression after 5 hours of induction.](image)

Cells were induced on: 0.5% Tween 80 (lane 1); 2% pristane (lane 2); 2% tetradecane (lane 3); 2% hexadecane & 2% pristane (lane 4); 2% octadecane & 2% pristane (lane 5); 2% stearic acid & 0.5% Tween 80 (lane 6); 2% docosane & 2% pristane (lane 7); 2% octacosane & 2% pristane (lane 8).

In a subsequent experiment a range of substrates were used to monitor the induction of all eight *Y. lipolytica* CYP52 genes. Glucose pre-grown cells were incubated in a medium containing no carbon source, as well as media containing either glucose, pristane, tetradecane, hexadecane, octadecane, or mixtures of docosane and pristane, octacosane and pristane, stearic acid and Tween 80, and behenic acid and Tween 80. Total cellular RNA was isolated after 5 hours of induction as was done for the range of substrates used to test YlALK1 induction levels (Fig. 4.6). Four membranes were prepared. Each membrane was probed with two of the eight *Y. lipolytica* CYP52 gene probes. Actin gene induction levels were monitored using membranes that showed the lowest CYP52 gene induction levels. Similarly, *Y. lipolytica* CYP52 gene induction levels were monitored using acid pre-grown cells incubated on the corresponding acids. Induction of six of the eight *Y. lipolytica* CYP52 genes can be seen for the glucose pre-grown and acid pre-grown cells (Fig. 4.7A & B).
Figure 4.7. Northern hybridization analysis of Y. lipolytica UOFS Y-0164 CYP52 gene expression after 5 hours. Cells were induced on: A: No C-source: Control (lane 1); 2% glucose (lane 2); 2% pristane (lane 3); 2% tetradecane (lane 4); 2% hexadecane (lane 5); 2% octadecane (lane 6); 2% docosane & 2% pristane (lane 7); 2% octacosane & 2% pristane (lane 8); 2% stearic acid & 0.5% Tween 80 (lane 9); 2% behenic acid & 0.5% Tween 80 (lane 10). B: 2% stearic acid & 0.5% Tween 80 (lane 1); 2% behenic acid & 0.5% Tween 80 (lane 2).
In general, YlALK1, YlALK2, YlALK3 and YlALK4 show stronger induction on the substrates tested than the remaining four genes. YlALK5 and YlALK6 appear to be induced to a lesser extent on the alkanes when comparing all eight genes (Fig. 4.7A). No mRNA transcripts were observed for either YlALK7 or YlALK8 (Fig. 4.7A & B), confirming the RT-PCR results. This could indicate that the conditions needed for induction were not met, or that either YlALK7 and/or YlALK8 might in fact be pseudogenes. It is difficult to compare the intensities of the signals obtained for YlALK1, YlALK2, YlALK3 and YlALK4 as the Actin signals obtained for the cells pre-grown on glucose were not the same for all of the samples tested (Fig. 4.7A). YlALK1 and YlALK4 appeared to be more strongly induced by the long-chain fatty acids using the glucose pre-grown cells. The same result was, however, not obtained when cells were pre-grown on the fatty acids (Fig. 4.7B).

The 824bp Actin probe was used as a quantitative control for the expression of the Y. lipolytica CYF52 genes. Due to the large number of substrates that were tested, it was not possible to induce all the yeast cells on all of the substrates of interest, at the same time. Despite the fact that growth conditions as well as RNA extraction procedures were kept as identical as possible, variations in the induction level of the Actin gene were evident (Fig. 4.7). This being the case, Actin signals observed for initial YlALK1 induction experiments (Fig. 4.6) mimic the results obtained for the CYF52 gene induction experiments (Fig. 4.7).

To monitor the expression of the eight Y. lipolytica CYF52 genes with time, cells were induced over a time-period of thirteen hours. The yeast cells grew most favourably on both octadecane and docosane (Fig. 4.3B), and as a result, total RNA samples were isolated at hours 5, 9 and 13 following induction on octadecane and docosane. Pristane was included in the induction study as it is needed as a co-solvent for docosane.

In general, YlALK1 shows stronger induction on the three substrates tested than the remaining seven P450ALKs (Fig. 4.8B). Induction patterns for YlALK2 to YlALK6 mimic those obtained in the five hour CYF52 induction study (Fig. 4.7A). The induction of YlALK1 through YlALK6 did not very significantly with time (Fig. 4.8B). No transcripts were observed for for either YlALK7 or YlALK8 for any of the substrates tested at any of the induction times (Fig. 4.8B).
Figure 4.8. A: Growth of *Y. lipolytica* UOFS Y-0164 on 2% pristane (●); 2% octadecane (○); and 2% docosane & 2% pristane (▲). B: Northern hybridization analysis of *Y. lipolytica* UOFS Y-0164 CYP52 gene expression with time. Cells were induced on 2% pristane for 5 hours (lane 1), 9 hours (lane 2), and 13 hours (lane 3); 2% octadecane for 5 (lane 4), 9 (lane 5), and 13 hours (lane 6); and 2% docosane & 2% pristane for 5 (lane 7), 9 (lane 8), and 13 hours (lane 9), respectively.
4.7 The relationship between the evolutionary distance and regulation of the *Y. lipolytica* UOFS Y-0164 CYP52 genes

To address the question of whether a relationship between evolutionary distance and gene induction by different substrates may be found, we created a phylogenetic tree using *Y. lipolytica*'s eight CYP52 gene amino acid sequences, and compared their evolutionary position with their substrate inducibilities (Fig. 4.9).

Three *YIALK* gene groups appear to exist. The first comprises the closely associated *YIALK1*, *YIALK2*, and the further removed *YIALK8*. *YIALK1* and *YIALK2* were both strongly induced on all the substrates tested. According to Iida *et al.* (2000), both *YIALK1* and *YIALK2* are needed for growth on longer-chain *n*-alkanes. The second cluster of *YIALK* genes forming a group comprises *YIALK3*, *YIALK5*, and *YIALK7*, while *YIALK4* forms the third group (Fig. 4.9). Groups two and three contain genes relatively weakly induced on both long-chain *n*-alkanes and fatty acids (Fig. 4.4, 4.6, 4.7 & 4.8). Here, *YIALK3* was induced preferentially on *n*-alkanes when compared to *YIALK4*'s fatty acid induction pattern (Fig. 4.7A). Individual groups therefore appear to have selective *n*-alkane/fatty acid induction patterns. Additional amino acid alignments of the eight *Y. lipolytica* CYP52 genes with the P450 genes present in *C. tropicalis*, *C. maliosa*, *C. apicola*, *C. albicans* and *B. megaterium* produced a phylogenetic tree with *YIALK8* forming an outgroup (data not shown). Identical results were produced by Iida *et al.* (2000).

---

![Phylogenetic relationship](image)

Figure 4.9. Phylogenetic relationship exhibited between the *Y. lipolytica* UOFS Y-0164 CYP52 gene amino acid sequences. The phylogenetic tree was established on the basis of a multiple sequence alignment done with the CLUSTALW program (see Section 3.2.10). Despite the application of various combinations of reasonable input parameters, the phylogenetic tree did not differ significantly.
Chapter Five

General discussion

Transcriptional control of the alkane hydroxylating cytochrome P450 monooxygenases is of interest due to the fact that these activities are important for the fermentative production of valuable alkane intermediates such as the dicarboxylic acids. Research focussing on the factors involved in the transcriptional regulation of the P450s found in alkane-assimilating yeasts has been hampered by the asexual and diploid genomes of the prominent Candida spp. Genetic studies using the yeast Yarrowia lipolytica are simplified in comparison to Candida spp. because of its haploid genome and sexual life cycle. As a result, genetic analysis of factors participating in P450 transcriptional control as well as studies focussing on the related enzymes, are made more approachable.

In 1998 Iida et al. described the detection of a cytochrome P450 in Y. lipolytica following growth on tetradecane with the subsequent cloning and characterization of the n-alkane-inducible gene YlALK1 (Iida et al., 1998). Screening of a Y. lipolytica CX-161 genomic library using an RT-PCR amplified probe identified a fragment encoding a 523 amino acid protein with a molecular mass of 60.2kDa. Expression of this gene showed induction on tetradecane and repression in the presence of glycerol and myristic acid. Subsequent screening of the same genomic library using degenerate PCR primers resulted in the isolation of an additional seven new genes, YlALK2-YlALK8 (Iida et al., 2000). Work done on other alkane-assimilating yeasts that contain CYP52 multigene families, have shown the individual genes to be regulated differently by different substrates and the individual enzymes to be substrate specific (Sanglard & Loper, 1989; Seghezzi et al., 1991 & 1992; Ohkuma et al., 1991a & 1991b; Schunck et al., 1991; Scheller et al., 1996; Zimmer et al., 1996; Lottermoser et al., 1996). Iida et al. (2000) showed this to be the case for the Y. lipolytica CYP52 multigene family. These authors investigated the individual CYP52 gene induction levels for short-chain n-alkanes ranging from C_{10}-C_{12}.

Generally there exist a number of approaches that can be used to study the involvement of individual genes, forming part of multigene families, in the utilization of various substrates. The first of these constitutes monitoring the inducible production of mRNA by way of Northern hybridizations as well as RT-PCR. Secondly, the heterologous expression of active enzyme systems in hosts such as S. cerevisiae, allows simplified characterizations of the individual enzymes of interest. Thirdly, the construction of gene disruptants can be used to investigate specific functions in vivo. This being the case, initial investigations largely focuss on mRNA induction levels for the sole reason that individual genes that form part of multigene families might very well be repressed under certain conditions. Consequently, such results are taken into consideration when performing either heterologous expression or in vivo disruptant studies. The aim of the present study was therefore to monitor the inducible production of Y. lipolytica CYP52 mRNA transcripts following exposure to very long-chain n-alkanes and fatty acids. The very long-chain n-alkanes and fatty acids investigated were docosane (C_{22}), octacosane...
(C_{28}) and behenic acid (C_{22}). For comparative purposes tetradecane (C_{14}), hexadecane (C_{16}), octadecane (C_{18}) and stearic acid (C_{18}) were also included in the study.

Over the past decade technologies for quantification of nucleic acids have been developed intensively thereby reflecting their importance in diagnostics and biomedical research. Methods available for this purpose include Northern hybridizations, quantitative competitive reverse transcription-PCR (QC RT-PCR), the RNase protection assay (RPA) and real time detection 5′-nuclease (RTDN) PCR. This being the case, such procedures have their advantages as well as their disadvantages. Hybridization methods have far less complex reaction kinetics when compared to PCR-based methods. PCR-based methods are more sensitive, but the use of PCR as a quantitative method is not as straightforward. In general, a quantitative RT-PCR method consists of four steps: i) generation of internal standards; ii) RT-PCR; iii) detection of products and iv) data analysis. The quantification of target nucleic acids using QC RT-PCR is calibrated by reference to an amount of a defined “competitor”. Such a “competitor” comprises a nucleic acid fragment that differs in size from the natural target of interest, but has a similar sequence to the amplified fragment. Due to the highly sensitive nature of PCRs, small differences in the physical properties of the target and its competitor can result in a significantly large difference in their total amplification efficiency. Consequently, the far less complex reaction kinetics of hybridization methods promoted the use of Northern hybridizations to monitor the inducible production of the mRNA transcripts of interest, rather than quantitation by means of RT-PCR.

Work done by Iida et al. (1998 & 2000) described the induction of the individual Y. lipolytica CYP52 genes on short-chain n-alkanes ranging from C_{10}-C_{16}. To date, the chain-length-dependent characterization of yeast cytochrome P450 monooxygenase forms has not exceeded C_{18} n-alkanes (Zimmer et al., 1996 & 1998; Ohkuma et al., 1998; Scheller et al., 1996). When planning and performing experiments to study the induction of the CYP52 genes in Y. lipolytica by very long-chain n-alkanes, several problems were encountered due to the use of the longer-chain substrates. Firstly, both the n-alkanes and fatty acids examined in this study were insoluble in water and therefore required a co-solvent. Pristane was used as the co-solvent for the longer-chain n-alkanes while Tween 80 was used to emulsify the long-chain fatty acids. For comparative purposes, growth conditions used by Iida et al. (1998) were implemented in this study which resulted in poor growth of the Y. lipolytica strain used here. Secondly, mRNA was isolated from the yeast cells prior to the exponential growth phase. As a result, the induction studies were performed on yeast cells not actively growing on the substrate of interest. The initial work done by Iida et al. (1998) used glucose-grown pre-cultures that resulted in long lag-phases prior to growth on the n-alkane of interest. Consequently, a large number of variables were introduced. These variables include successive pre-cultures on consecutive days due to the number of substrates examined, the hydrophobic nature of the cells in the presence of long-chain fatty acids as well as co-solvent complications.

Initial attempts to obtain high-quality RNA preparations from Y. lipolytica cells using the method described by De Winde et al. (1996) failed to produce results for both the RT-PCRs and the Northern hybridizations. The latter method was optimized for S. cerevisiae,
but, it is considerably more difficult to isolate high-quality RNA from *Y. lipolytica* (Mauersberger *et al.*, 1996). A second method described by Domdey *et al.* (1984) produced higher-quality RNA when compared to the De Winde method, and was subsequently used for the induction studies.

The identification of the individual *Y. lipolytica CYP52* genes for the induction studies required eight, gene-specific, *CYP52* probes. Eight gene-specific primers were designed from the non-conserved 5'-regions of the eight *Y. lipolytica CYP52* genes, which were then used in a PCR-based reaction. Examination of the probes used by Iida *et al.* (2000) for the characterization of the *Y. lipolytica CYP52* multigene family showed that the individual probes were designed identically to ours. Both sets of probes were designed from the 5'-region of the genes and the corresponding sizes of the eight gene probes were approximately the same. *Y. lipolytica CYP52* nucleic acid sequence alignments showed that the remaining regions of the genes were far too conserved to prepare eight gene-specific probes. Although the two sets of probes were prepared from the same *Y. lipolytica CYP52* nucleic acid region, individual gene nucleic acid sequence alignments showed that the different strains used in the two studies had variable nucleic acid sequences.

Initially, the induction of the eight *Y. lipolytica CYP52* genes was monitored using RT-PCR reactions. RT-PCR reaction conditions were optimized such that one reaction programme could be used for all eight genes on both the longer-chain n-alkanes and fatty acids. This was the first time that the RT-PCR reaction was used to monitor the induction of *CYP52* genes in any alkane-assimilating yeast. Unfortunately, the individual *CYP52* gene induction levels could not be quantitated with a normal RT-PCR reaction due to the complex reaction kinetics. The optimization of the initial PCR reaction conditions allowed one reaction programme for the amplification of all eight *CYP52* gene products, although the individual band intensities were not all identical. As a result, differences in the individual band intensities for the RT-PCR reactions cannot be attributed to varying gene induction levels. Despite the fact that the RT-PCR reactions could not be used to quantitatively compare any induction levels, the reactions clearly showed which genes were induced and which were not. Here, six of the eight *Y. lipolytica CYP52* genes were induced on most of the substrates tested. No mRNA transcripts were observed for either *YIALK7* or *YIALK8* on any of the substrates tested. This result was confirmed by Iida *et al.* (2000) using n-alkanes ranging from C_{10} to C_{14}, although they speculated that *YIALK7* and *YIALK8* might be involved in the utilization of longer-chain n-alkanes. Attempts to quantify *Y. lipolytica CYP52* gene expression levels using real-time RT-PCR yielded poor results (data not shown). Primer-dimer formation significantly hampered quantification of the *CYP52* and Actin gene expression using the LightCycler System (Roche). Due to poor primer design and time constraints the quantification of the individual *Y. lipolytica CYP52* genes was not possible.

Northern hybridizations were used in an attempt to quantify the *Y. lipolytica CYP52* gene induction levels following exposure to the longer-chain n-alkanes and fatty acids of interest. Conditions that affected the quantitative ability of the Northern hybridizations were, among others, comparing the induction levels of a large number of genes on an
even larger number of substrates. Irregularities in the expression of the Actin gene, used as a quantitative control, further complicated the comparison of the individual CYP52 induction specificities. Although the Northern hybridization experiments were not quantitatively conclusive, the results obtained corroborate the results obtained using the RT-PCR reactions as well as those obtained by Iida et al. (1998 & 2000). As was seen for the RT-PCR reactions, neither YIALK7 nor YIALK8 appear to be induced on any of the substrates tested at any of the induction times. Generally, YIALK1 through YIALK6 did not show chain-length-dependent induction specificities although YIALK1 through YIALK4 had stronger induction signals when compared to YIALK5 and YIALK6. These results are comparable to those obtained by Iida et al. (2000) in which YIALK1 through YIALK4 were preferentially induced by shorter-chain n-alkanes while YIALK1 and YIALK2 were found to be essential for hexadecane assimilation. Interestingly, only YIALK2 was observed to be weakly induced on myristic acid (Iida et al., 2000) whilst we found YIALK1, YIALK2 and YIALK4 to be strongly induced on both stearic and behenic acid. This could possibly be attributed to the fact that Tween 80 was included in the induction medium whilst Iida et al. (2000) made no mention of the use of Tween 80 as an emulsifier. Additionally, cells pre-grown on a complex fatty acid medium produced different CYP52 induction specificities when compared to the induction specificities of cells pre-grown on glucose.

The characterization of the Y. lipolytica CYP52 multigene family in terms of gene induction levels can be compared to the CYP52 multigene families belonging to C. maltosa and C. tropicalis. For example, the initial investigations describing the C. tropicalis CYP52 gene induction specificities showed six of the seven genes to be switched on in the presence of n-alkanes. The seventh gene, CYP52A7, was initially thought to be a pseudogene although successful expression of CYP52A7 in S. cerevisiae proved otherwise (Seghezzi et al., 1992). Similarly, the enzymatic characterization of the C. maltosa cytochromes P450 has shown the yeast to contain four major n-alkane-inducible P450 forms (Zimmer et al., 1996), as well as two fatty acid-inducible P450s (Zimmer et al., 1998). Consequently, six of the eight Y. lipolytica CYP52 genes found to be n-alkane-inducible, need to be heterologously expressed in S. cerevisiae. Doing so will establish whether the individual genes produce active enzymes involved in n-alkane-assimilation. Although YIALK7 and YIALK8 were not found to be induced by n-alkanes or fatty acids, heterologous expression of the individual P450s in S. cerevisiae will determine whether or not they code for active enzymes as was done for CYP52A7.

The induction of both yeast peroxisomal β-oxidation and CYP52 genes in the presence of fatty acids indicates that the individual genes might in fact be regulated in a similar fashion. Examination of the eight Y. lipolytica CYP52 promoter regions revealed the existence of specific upstream regulating elements (UREs). The promoter regions for YIALK1, YIALK2, YIALK4, and YIALK7 contained an oleate response element (ORE), CGGNNNTNA, responsible for gene activation in the presence of fatty acids (Filipits et al., 1993). Interestingly, the Y. lipolytica multifunctional enzyme type 2 (MEF2) gene that forms part of the genes that contribute to β-oxidation in yeast, contains a putative upstream oleic acid-response element (Smith et al., 2000), as do those in S. cerevisiae and C. tropicalis (Sloots et al., 1991; Filipits et al., 1993). Additional analysis of the
promoter region for YIALK4 revealed the sequence MRGAAAT, responsible for overcoming glucose repression (Sloots et al., 1991). Such glucose repression systems have been identified in the C. tropicalis and C. maltosa CYP52 multigene families (Seghezzi et al., 1992; Ohkuma et al., 1998). Incidentally, the Y. lipolytica CYP52 multigene family is not repressed by glucose but by glycerol.

In addition to the information regarding the Y. lipolytica CYP52 n-alkane and fatty acid gene inducibilities, we attempted to gain more insight into the P450 structure-function relationship in an evolutionary context. Phylogenetic analysis of the individual Y. lipolytica CYP52 genes compares well with the phylogenetic tree constructed by Iida et al. (2000). Three groups are evident in both. It is suggested that Y/ALK1 and Y/ALK2, that form the first group in both trees, belong to the early branched group among the Y. lipolytica P450ALKs (Iida et al., 2000). Here, the furthest removed diverse Y. lipolytica P450ALKs seen in both trees, namely Y/ALK4 and Y/ALK5, are believed to have evolved through gene duplication (Iida et al., 2000).

In summary, the results obtained by Iida et al. (1998 & 2000), as well as those obtained in the present study, demonstrate that Y. lipolytica contains n-alkane-inducible P450 monoxygenases that exhibit different functions in the metabolism of long-chain hydrocarbons and fatty acids. Further analysis of the individual Y. lipolytica CYP52 genes in terms of their properties and expression should reveal the mechanisms of genetic control of the P450ALKs in this yeast. Such knowledge may provide a suitable basis for future studies on P450 structure-function relationships as well as for the biotechnological application of the Y. lipolytica P450 forms.
References


83


Summary

A large number of monooxygenases contain a haem protein containing cytochrome P450. These monooxygenases enzymatically catalyze dioxygen activation at the cytochrome P450 haem protein. The cytochrome-P450-dependent monooxygenases are involved in many steps of the biosynthesis and the degradation of compounds such as steroids, fatty acids, prostaglandins, leukotrienes, and n-alkanes. The first enzymatic step of hydrocarbon assimilation is the terminal hydroxylation of the n-alkane by a cytochrome P450 enzyme system. The enzyme system bound to the endoplasmic reticulum (ER), consists of an NADPH-cytochrome P450 reductase involved in transferring electrons (Govindaraj & Poulos, 1997; Nelson & Strobel, 1988), and a cytochrome P450 acting as hydroxylase in a typical monooxygenase reaction.

The cytochrome P450 monooxygenase (CYP52) multigene family involved in the terminal hydroxylation of n-alkanes and fatty acids has been characterized in yeast species such as Candida maltosa, Candida tropicalis and Candida apicola (Lottermoser et al., 1996; Seghezzi et al., 1992; Zimmer et al., 1996). These include eight genes in C. maltosa, seven in C. tropicalis, and two in C. apicola. Yarrowia lipolytica, originally classified as a Candida, uses few sugars (mainly glucose) as a carbon source. This yeast readily assimilates organic acids, various polyalcohols, and normal paraffins. The first of eight Y. lipolytica cytochrome P450 monooxygenase characterized was Y/ALK1 (Iida et al., 1998). Subsequent characterization of the remaining seven Y. lipolytica CYP52 genes showed Y/ALK1 and Y/ALK2 to be the major P450 forms involved in assimilation of shorter-chain n-alkanes (C10-C16), while the remaining Y/ALK genes (Y/ALK3 through Y/ALK8), did not appear to be significantly involved in C10-C16 assimilation (Iida et al., 2000).

The induction of the eight Y. lipolytica CYP52 genes on longer-chain n-alkanes (C18-C28) and long-chain fatty acids was investigated using gene-specific RT-PCR reactions as well as Northern hybridizations. A PCR-based approach was used to prepare eight gene-specific Y. lipolytica CYP52 probes. The probes designed for this study were the same as those used by Iida et al. (2000). Complications that hampered the study included: (i) the insolubility of the very long-chain n-alkanes and fatty acids in water, which necessitated the use of co-solvents such as pristane and Tween 80; (ii) the poor growth of the selected yeast strain under the conditions used by Iida et al. (2000) for induction of the genes; (iii) the large number of genes investigated; (iv) the logistical problems associated with comparing a large number of genes under so many growth conditions, and (v) the difficulty in isolating RNA from Y. lipolytica.

Induction of the Y. lipolytica CYP52 genes were studied in the presence of glucose, tetradecane, hexadecane, octadecane, mixtures of docosane and pristane, octacosane and pristane, stearic acid and Tween 80 and behenic acid and Tween 80. Negative controls constituted cultures without any substrate as well as cultures containing only either pristane or Tween 80. The first two Y. lipolytica CYP52 genes investigated, Y/ALK1 and Y/ALK2, showed preferential induction on all the substrates tested. Genes Y/ALK3 through Y/ALK6 showed relatively weaker induction on the substrates tested when
compared to YIALK1 and YIALK2. No mRNA transcripts were observed for either
YIALK7 or YIALK8 on any of the substrates tested at any of the induction times. These
results were not only repeated for both the RT-PCR and Northern hybridization
experiments, but were confirmed for the shorter-chain n-alkanes by Iida et al. (2000).
Variations in the levels of Y. lipolytica CYP52 gene transcripts were observed when using
glucose pre-grown or acid pre-grown yeast cells for the induction experiments.

Phylogenetic analysis of the individual Y. lipolytica CYP52 genes suggested that this
multigene family evolved and diverged after branching off from the ancestral P450ALK
gene. YIALK1 and YIALK2, which were most prominent in the induction studies,
grouped together in the phylogenetic analysis. Comparison of the Y. lipolytica CYP52 and
peroxisomal β-oxidation MFE2 gene promoter regions revealed oleic acid-response
elements involved in activation by fatty acids. Elucidating the transcriptional activating
sequences present in the β-oxidation and CYP52 enzyme systems could lead to an
understanding of the regulation of enzymes that contribute to the terminal- and β-
oxidation reactions occurring within these n-alkane-assimilating yeasts.
Opsomming

'n Groot aantal monoëksigenase bevat heemproteïne wat sitochroom P450 bevat. Hierdie monoëksigenase kataliseer dioksigeenaktivering by die sitochroom P450 heemproteïne ensimaties. Die sitochroom-P450-afhanklike monoëksigenase is betrokke by die vele stappe van biosintese en die afbreking van verbindinge soos steroïde, vetsure, prostaglandiene, leukotrieens, en n-alkane. Die eerste ensimatiese stap van koolwaterstofassimilering is die terminale hidroksilering van die n-alkane deur 'n sitochroom P450 ensiemstel. Die ensiemstel wat aan die endoplasmatiëse retikulum (ER) verbind is, bestaan uit 'n NADPH-sitochroom P450 reduktase wat by die oordra van elektrone betrokke is (Govindaraj & Poulos, 1997; Nelson & Strobeol, 1988), en 'n sitochroom P450 wat as hidroksilase in 'n tipiese monoëksigenase reaksie te werk gaan.

Die sitochroom P450 monoëksigenase (CYP52) multigeefamilie wat by die terminale hidroksilering van n-alkane en vetsure betrokke is, is gekarakteriseer in gissoorte soos Candida maltosa, Candida tropicalis en Candida apicola (Lottermoser et al., 1996; Seghezzi et al., 1992; Zimmer et al., 1996). Hierdie bevat agt gene in C. maltosa, sewe in C. tropicalis, en twee in C. apicola. Yarrowia lipolytica, oorspronklik geklassifiseer as 'n Candida, gebruik enkele soorte suiker (hoofsaaklik glukose) as 'n koolstofbron. Hierdie gis assimileer maklik organiese sure, verskeie tipes poli-alkohol, en gewone paraffiensoorte. Die eerste van agt Y. lipolytica sitochroom P450 monoëksigenase wat gekarakteriseer is, was YIALK1 (Iida et al., 1998). Daaropvolgende karakterisering van die oorblywende sewe Y. lipolytica CYP52 gene het YIALK1 en YIALK2 aangetoon as die hoofvorme van P450 wat by die assimilering van korter-ketting n-alkane (C\textsubscript{16}-C\textsubscript{18}) betrokke is, terwyl die oorblywende YIALK gene (van YIALK3 tot YIALK8), nie gebyk beduidend betrokke te gewees het nie by C\textsubscript{10}-C\textsubscript{16} assimilering (Iida et al., 2000).

Die indusering van die agt Y. lipolytica CYP52 gene op langer-ketting n-alkane (C\textsubscript{18}-C\textsubscript{28}) en langketting vetsure is ondersoek deur die gebruik van geen-spesifieke RT-PCR reaksies asook Northern hibridiserings. 'n PCR-gebaseerde benadering is gebruik om agt geen-spesifieke Y. lipolytica CYP52 ondersoek voor te berei. Die ondersoek is deur hierdie studie ontwerp is, was dieselfde soos wat deur lida et al. (2000) gebruik is. Komplikasies wat hierdie studie bemoeilik het, het die volgende ingesluit: (i) die onoplosbaarheid van die baie lang-ketting n-alkane en vetsure in water, wat die gebruik van medeoplossers soos pristaan en Tween 80 genoodsaak het; (ii) die swak groei van die geselekteerde gisstam onder die toestande wat deur lida et al. (2000) vir die indusering van gene gebruik is; (iii) die groot getal gene wat ondersoek is; (iv) die logistiese probleme wat geassosieer word met die vergelyking van 'n groot getal gene onder veelvuldige groeitoeands, en (v) die moeisaamheid om RNA van Y. lipolytica te isolateer.

Die indusering van die Y. lipolytica CYP52 gene is bestudeer in die teenwoordigheid van glukose, tetradekaan, heksadekaan, oktadekaan, dokosaan met pristaan, oktakosaan met pristaan, steariensuur met Tween 80 en behensuur met Tween 80. Negatiewe kontrole het kulture saamgestel sonder enige substraat asook kulture wat slegs pristaan of Tween 80 bevat. Die eerste twee Y. lipolytica CYP52 gene wat ondersoek is, nl. YIALK1 en YIALK2, het 'n voorkeurindusering getoon op alle substrate wat getoets is. Gene YIALK3
tot by YIALK6 het ‘n relatief-swakker indusering getoon op die substrate wat getoets is wanneer dit met YIALK1 en YIALK2 vergelyk is. Geen mRNA transkripte is waargeneem hetsy vir YIALK7 of YIALK8 nie of enige van die substrate wat op enige van die induseringstye getoets is nie. Hierdie resultate is nie net vir beide die RT-PCR en Northern hibridiseringseksperimmente herhaal nie, maar is vir die korter-ketting n-alkane bevestig deur Iida et al. (2000). Variasies in die vlakke van Y. lipolytica CYP52 geentranskripte is waargeneem wanneer glukose-voorafgekweekte of suurvoorafgekweekte gisselle vir induseringseksperimmente gebruik is.

Filogenetiese ontleding van die individuele Y. lipolytica CYP52 gene het daarop gedui dat hierdie multigenefamilie ontwikkel en uiteengelopen het nadat dit van die voorouerlike P450ALK geen vertak het. YIALK1 en YIALK2, wat baie prominent was in die induseringstudies, het saamgegroepeer in die filogenetiese ontleding. Vergelyking van die Y. lipolytica CYP52 en peroksisomale β-oksidasie MFE2 geenbevorderingsgebiede, het oleïensuur-responseelemente betrokke by die aktivering van vetsure, blootgelê. Die opheldering van die transkripsionele aktiveringsreekse teenwoordig in die β-oksidasie en CYP52 ensiemstelsels, kan lei tot ‘n begrip van die regulering van ensieme wat bydra tot die termaal- en β-oksidasiereaksies wat binne hierdie n-alkaanassimileringsgiste plaasvind.