HETEROLOGOUS EXPRESSION OF CYTOCHROME P450 MONOOXYGENASES BY THE YEAST *YARROWIA LIPOLYTICA*

BY

CHRISPIAN WILLIAM THERON

SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

MAGISTER SCIENTIA

IN THE

DEPARTMENT OF MICROBIAL, BIOCHEMICAL & FOOD BIOTECHNOLOGY
FACULTY OF NATURAL AND AGRICULTURAL SCIENCES
UNIVERSITY OF THE FREE STATE
BLOEMFONTEIN 9300
SOUTH AFRICA

MAY 2007

SUPERVISOR: PROF. M.S. SMIT
CO-SUPERVISOR: PROF. J. ALBERTYN

“Insanity: doing the same thing over and over again and expecting different results."

Albert Einstein

"Reality is merely an illusion, albeit a persistent one"

Albert Einstein

"Not all who wander are lost"

JRR Tolkien

Acknowledgements

I would like to express my sincerest gratitude towards the following people:

The Almighty Heavenly Father, for this gift of life.

Prof. M.S. Smit for her invaluable knowledge and endless patience throughout the course of this project. Thank you for your tireless effort!

Prof. J.A. Albertyn, for guiding me through the molecular conundrums of this project.

Dr. G.O. Obiero, for all of the valuable guidance about science, and about life.

Dr. M.E. Setati, for all of the interesting and thought-provoking conversations.

Mr. Piet Botes for his technical assistance with chemical analysis.

DST-NRF Centre of Excellence in Catalysis (C*Change) for the financial support of this project.

My friends from the **Biotransformations research group.**

The Molecular Biology research group, special thanks to **Dr Khajamohiddin Syed.**

Fellow friends and colleagues in the Department of Biotechnology, UFS.

Family and friends for their love, support and encouragement. Special thanks to **Julian and Maruska** for their support.

Finally, but certainly not least, **Reinhild Teise** for her moral support, caring, love and patience that she showed me during the rough times of this project.

Table of contents

Chapter 1: Introduction

1.1. Motivation for present study

Biocatalysis is the use of biological systems, whole cells or (partially) isolated enzymes, to enhance chemical reactions. Of various types of enzymes used, alkane hydroxylases have generated particular interest due to their abilities to perform difficult chemistry. Their ability to facilitate oxygenation of non-activated carbons is intriguing, as this cannot be replicated to nearly the same efficiency using synthetic chemistry. The enzymatic processes have further advantages over traditional chemical syntheses in their excellent chemo-, regio, and stereo-selectivities (Glieder *et al*, 2002).

Additional advantages using biocatalysts include less extreme reaction conditions, which in turn results in increased safety of processes, energy savings and lower operating costs, and reduced environmental damage, contributing to 'green chemistry' processes (Ayala and Torres, 2004; Glieder *et al* 2002).

These factors make the use of biochemical means for improvement of current pure chemical syntheses a promising field of research indeed and these biological processes can find application in various industrial processes where cheap hydrocarbon substrates, usually wasteful by-products of chemical processes, can be converted to valuable products via hydroxylation. The petrochemical industry is an example, where inexpensive petroleum compounds can be converted to higher value products via C-H bond activation. Industrial applications of these enzymes also reach into the agrochemical, food and pharmaceutical industries (Ayala and Torres, 2004).

Different ubiquitous enzymes involved in aerobic activation of carbon-hydrogen bonds have been identified as membrane bound diiron alkane hydroxylases,

membrane-bound copper- (and possibly iron-) containing methane monooxygenases, soluble methane monooxygenases and cytochrome P450 enzymes (van Beilen and Funhoff, 2005; van Beilen and Funhoff, 2007).

These enzymes tend to be part of multi-component systems, as they generally receive reducing equivalents via reductase elements. This complicates the use of isolated enzymes for biocatalytic applications. Within the P450 enzymes though, a catalytically self-sufficient class has been identified, which consist of single protein components dependant on NAD(P)H (Narhi and Fulco, 1986). The single component nature of these enzymes increases their allure as biocatalysts, as the electron transfer intermediates are eliminated, speeding up the reactions, while simultaneously avoiding the problems encountered with multi-component systems. Mutants of this enzyme have been generated which have substrate specificity towards alkanes, and have proven to be the most efficient hydroxylases for short to medium chain length alkanes discovered to date (Glieder *et al*, 2002; Peters *et al*, 2003).

The cytochrome P450 enzymes appear to be the most useful alkane hydroxylases, and are certainly the best characterized systems. As an enzyme superfamily, they are the most diverse in terms of substrate specificities, electron transfer systems and sub-cellular localization. The catalytic heme has a very high affinity for oxygen, and hence efficiently uses molecular oxygen and reducing equivalents to hydroxylate aliphatic hydrocarbon substrates (Werck-Reichart and Feyereisen, 2000; Van Beilen and Funhoff, 2005; Van Beilen and Funhoff, 2007).

Although cytochrome P450 enzymes are commonly involved in bioremediation processes, they are also widely used for biotransformations of low value substrates to valuable products. One such example is the use of limonene, a waste product of the orange juice industry obtained from orange peels, which was converted to perillyl alcohol, a compound of value to the flavor and fragrance industry, as well as being a potential anti-cancer drug (Van Beilen *et al*, 2005).

1.2. Introduction to cytochrome P450 monooxygenases

Cytochrome P450 proteins are a ubiquitous superfamily of proteins. They are named after the absorption band produced at 450 nm when carbon monoxide is bound to the reduced heme of the active site. The nomenclature for these enzymes was recommended by a nomenclature committee presided over by Dr David Nelson (http://drnelson.utmem.edu/CytochromeP450.html). Following the root CYP, they are grouped into families (denoted by a number), then sub-families (denoted by a letter). Members of the same family share at least 40% amino acid identity, while members of the same sub-family share at least 55% amino acid identity (Kelly *et al*, 2003; Werck-Reichhart and Feyereisen, 2000).

There is a conserved cysteine residue serving as a fifth ligand to the heme iron (figure 1.1). The general reaction catalyzed by the P450 monooxygenases is the hydroxylation of a carbon-hydrogen bond via the reductive cleavage of molecular oxygen to water (figures 1.2 and 1.3). The strong affinity of the heme for oxygen is the key to the ability of these enzymes to use molecular oxygen as oxidant for hydroxylations, and plays a role in their remarkable ability to selectively activate chemically inactive carbons via introduction of oxygen into carbon-hydrogen bonds. Hydroxylations are the most common reactions catalyzed by these enzymes, although epoxidations, sulfoxidations and dealkylations have also been described (Urlacher *et al*, 2004; Kelly *et al*, 2003).

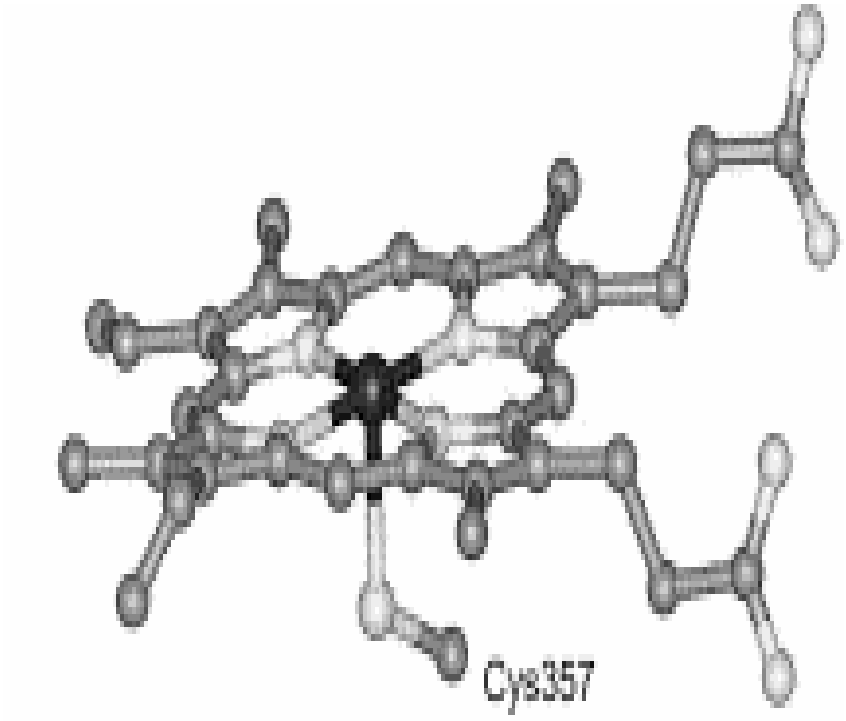
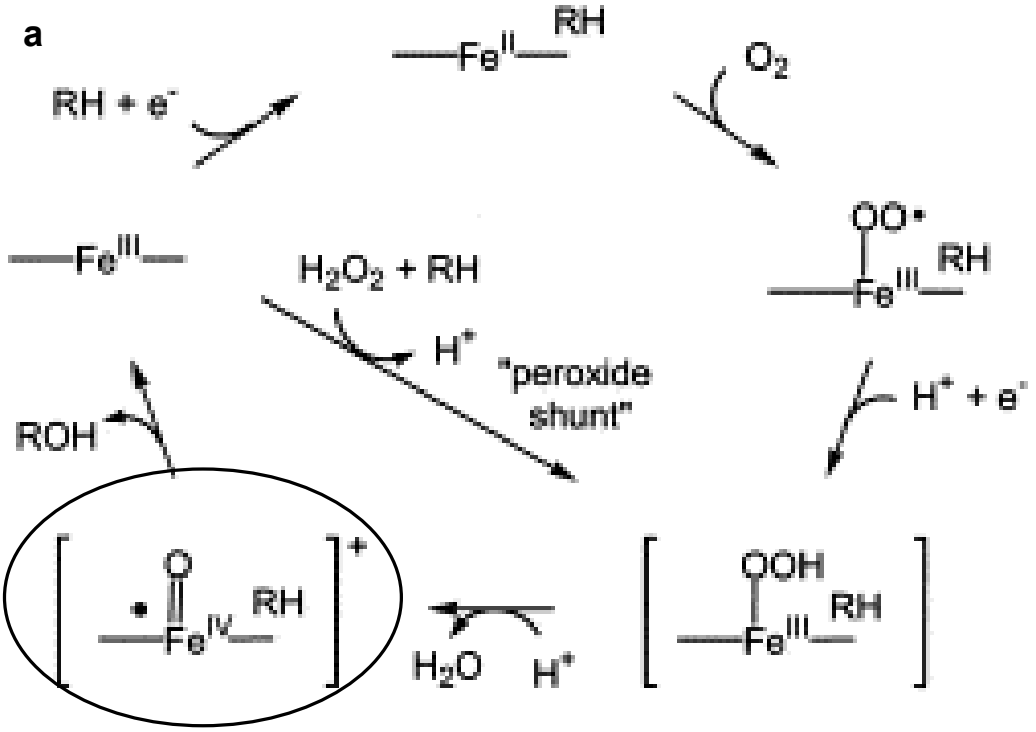


Figure 1.1: Model of P450 active site. The conserved cysteine residue is shown with the sulfur group from cysteine bound to the heme iron as its fifth ligand (from Costas *et al*, 2000).

$$R\text{-}H + O_2 + 2H^+ + 2e^- \xrightarrow{P450} R\text{-}OH + H_2O$$

Figure 1.2: General hydroxylation reaction catalyzed by cytochrome P450s (from Urlacher *et al*, 2004).

a


RH + e⁻
RH
O2
H2O2 + RH
H+
"peroxide shunt"
H+ + e⁻
ROH
H2O
H+

b

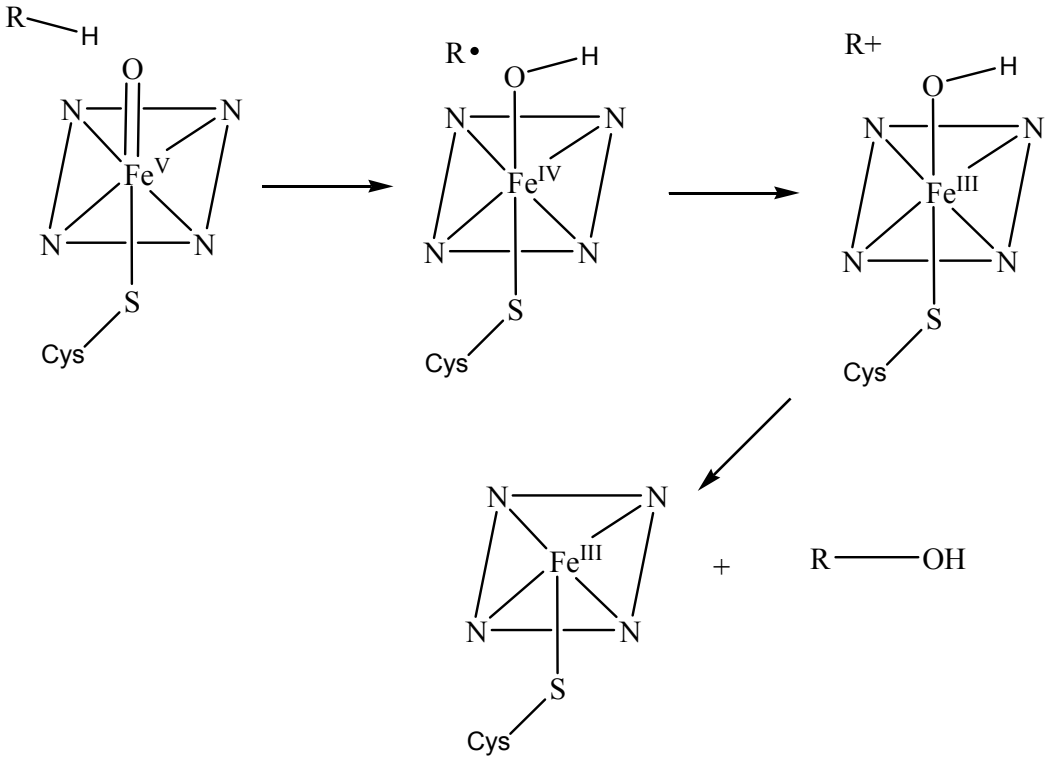

R
H
R•
R+
O
N
Fe^V
Fe^IV
Fe^III
S
Cys
R
OH
+

Figure 1.3: Proposed reaction mechanism for cytochrome P450. (a) The reaction starts by substrate binding to the enzyme while in the Fe^{III} state, triggering the first electron initially donated by NAD(P)H for reduction of Fe^{III}. Next is the binding of dioxygen (O_2), followed by the transfer of the second electron from NAD(P)H and one proton, generating a peroxide containing intermediate, Fe^{III}-OOH. A second proton is donated, forming water which leaves the reaction, and subsequently the reactive species Fe^{V}=O, or $[(Por\bullet)Fe^{IV}=O]^{+}$ (encircled) is produced, which is responsible for substrate oxygenation. Thereafter, the hydroxylated product departs, and the enzyme moves back to its original state for substrate binding, Fe^{III}. It is also suggested that, at least for some P450s, the Fe^{III} form can react with hydrogen peroxide, forming the reactive species in fewer steps in a process referred to as a peroxide shunt (from Costas *et al*, 2000). A suggestion for the detailed action of the reactive intermediate is the rebound mechanism (b). The theory behind this is that the reactive species abstracts a proton from the hydrocarbon substrate, leaving a substrate radical intermediate. In the meantime the active site has been reduced to Fe^{IV}. The radical "rebound" then occurs, where the radical substrate species reclaims the proton along with its associated oxygen, releasing the hydroxylated substrate and regenerating the Fe^{III} species for substrate binding (adapted from Moe *et al*, 2004).

Cytochrome P450s are oxidoreductase proteins, and thus require reducing equivalents for the reduction half of the reactions, supplied by NAD(P)H. The means by which the electrons from NAD(P)H reach the catalytic site, decides the classification of the type of P450. One such classification system discriminates between four classes. Class I proteins require two additional electron transfer proteins, as they receive their electrons from the NAD(P)H via an iron sulfur redoxin (ferredoxin), and an FAD-containing ferredoxin reductase. Proteins from class II have their own specific cytochrome P450 reductase containing FAD and FMN. Class III enzymes are known as self-sufficient P450s, as they have their own specific cytochrome P450 reductase domains as in class II, but now fused to

the hydroxylase in a single polypeptide. They therefore receive their electrons directly from the NAD(P)H. Class IV P450s receive their electrons from an iron-sulfur and FMN-containing reductase resembling those found in the phthalate family of proteins, which is also fused to the hydroxylase component in a single polypeptide (Roberts *et al*, 2002; De Mot and Parret, 2002; Werck-Reichhart and Feyereisen, 2000). These different classes are illustrated in figure 1.4.

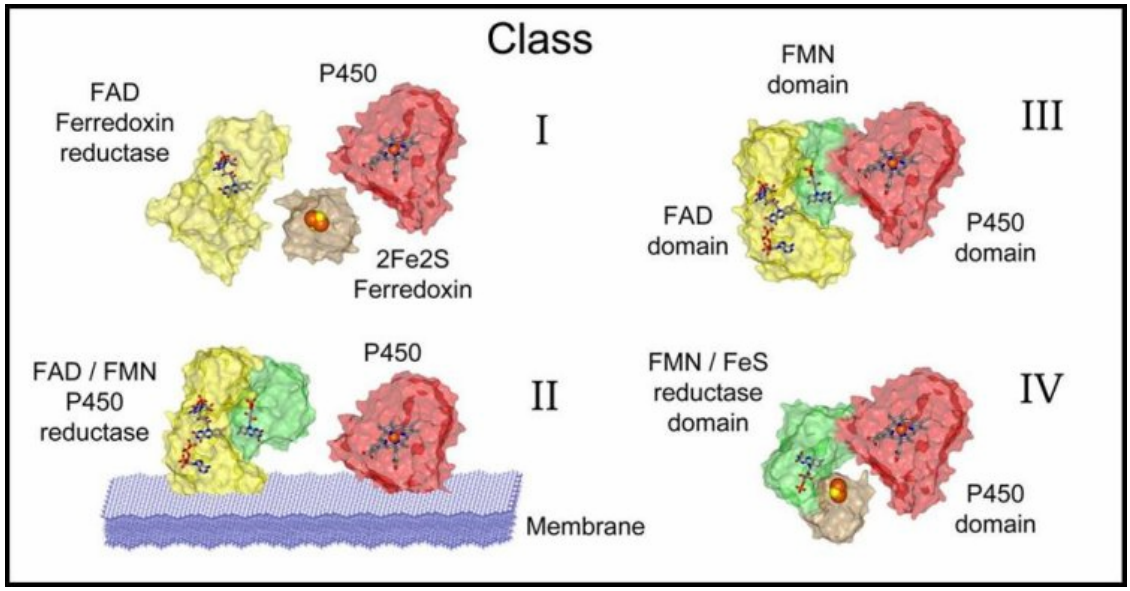


Figure 1.4: Different classes of cytochrome P450s. (a) An explicit illustration of the different classes and the intermediates involved in electron transport (from http://www.chem.ed.ac.uk/chapman/p450.html).

Recently, a more meticulous classification system has been described, separating P450s into 10 different classes based on their source, components of their electron transfer chain (including initial electron supply), and localization of P450 in the cell (Hannemann *et al*, 2007). The four-class system is sufficient for our purposes at this time though.

CYP450s are generally involved in metabolism of hydrophobic substrates in prokaryotes and simpler eukaryotes, production of metabolites in plants and

microbes, from which a great number of pharmaceuticals have been discovered, and in drug metabolism and toxicity in mammals. Common applications of these enzymes have been biotransformations of inexpensive substrates to more valuable products and bioremediation of especially organic environmental pollutants (van Beilen and Funhoff, 2007).

In bioremediation, P450 enzymes catalyze the initial activation steps of contaminants, such as dehalogenation of haloaromatic compounds or hydroxylation to increase the hydrophilicity of the compounds. Thereafter the compounds can be taken up by other mineralization pathways in especially bacterial cells (Wong *et al*, 1998).

Biotransformations focus mainly on replacing or improving historically used chemical syntheses. The high regio-, enantio- and stereo-selectivity of CYP450s, as well as the mild reaction conditions required by them, are among the characteristics favoring the use of these proteins over chemical syntheses. The greatest drawback of using CYP450s however, is their unavoidable requirement for continuous supply of reducing equivalents. The NAD(P)H involved may inhibit P450s in the absence of substrate, and is also expensive. While these problems are overcome when whole cells are used, only limited success has been achieved thus far using this strategy. Problems using whole cells include oxygen mass transfer, cell toxicity of some substrates and uptake of some hydrophobic substrates (Schwaneberg *et al*, 2000; Van Beilen and Funhoff, 2005; Van Beilen and Funhoff, 2007).

Self-sufficient P450s

An ideal biocatalyst for industrial applications would be soluble, self-sufficient / independent of NAD(P)H, stable under a range of conditions, and be highly regio-, enantio-, and stereoselective (Van Beilen *et al*, 2005).

A soluble P450 protein from *Bacillus megaterium* was identified and found to be responsible for the observed sub-terminal mono-hydroxylation of fatty acids in the presence of NADPH and O_2 catalyzed by cell free extracts of this organism. Furthermore, the enzyme also catalyzed the hydroxylation and epoxidation of unsaturated fatty acids, in ratios dependant on the position of the double bond (Miura and Fulco, 1974; Ruettinger and Fulco, 1981).

This monooxygenase was initially termed P450 BM-3, as it was the third P450 isolated from barbitute-induced *Bacillus megaterium* (Nahri and Fulco, 1986). It was later titled CYP102A1 according to the previously mentioned P450 nomenclature system. This enzyme was purified and further characterized. It was found that when NADPH and O_2 were present, substrate hydroxylation occurred without the requirement of any other protein component. The enzyme also showed strong reduction of cytochrome *c* in the presence of NADPH, indicating the presence of a reductase component. Through limited trypsin proteolysis they were able to identify the two separate domains, which retain their original half reactions after separation. After proteolysis, *in vitro* fatty acid monooxygenase activity could not be reconstituted, and the separate domains show no affinity for each other. The protein was confirmed to be a P450 domain fused to a specific eukaryotic-type reductase domain in a single polypeptide, making it essentially self-sufficient (Narhi and Fulco, 1987; Narhi and Fulco, 1986).

Wild-type P450 BM-3 has a specificity for C_{12}-C_{22} fatty acids, and was found to be the fastest P450 monooxygenase known, with catalytic rates exceeding 5000 min^{-1} for preferred substrates, for example arachidonic acid (Noble *et al*, 1999). The enzyme was cloned under its wild-type promoter into *E. coli*, resulting in a protein that was, when isolated, identical to the wild-type enzyme in all aspects (Narhi *et al*, 1988). During heterologous expression of the P450 BM-3 from *B. megaterium* in *E. coli*, cell-free extracts had approximately 24-fold higher activity than whole cell activity. The differences were likely due to ineffective uptake of

hydrophobic substrates into the *E. coli* cells, and the lack of an effective regeneration system for reduced co-factors in *E. coli* (Schneider *et al*, 1999).

Since the initial interest in P450 BM-3, there has been an astounding amount of research dedicated to these and related enzymes, as they seem to fulfill most of the requirements of an optimal biocatalyst. A range of mutant enzymes have further increased the substrate specificity and catalytic rates of these enzymes. Among these, P450 BM-3 has been mutated into a highly active alkane hydroxylase, with a higher turnover than any other reported hydroxylase for shorter chain length alkanes (C3-C8). Variations of these mutants also efficiently hydroxylate longer chain alkanes with improved selectivity. The alkane hydroxylating mutants also showed higher fatty acid hydroxylase activities than the wild-type enzyme (Glieder *et al*, 2002; Peters *et al*, 2003). Further mutations led to an efficient alkene epoxidase, which also hydroxylated benzene, forming phenol (Farinas *et al*, 2004). Another mutant of CYP450 BM-3 was generated to be dependant on NADH rather than NADPH for reducing equivalents, allowing potential application of a cofactor regeneration system involving another enzymatic reaction (Maurer *et al*, 2005).

CYP102A2 and CYP102A3 have since been identified in *B. subtilis*, showing 76% and 77% amino acid similarity respectively to CYP102A1 of *B. megaterium*. Similar results were obtained for their wild-type enzyme activities compared to CYP102A1, as well as for mutations of these enzymes which corresponded to some mutations of CYP102A1 (Lentz *et al*, 2004).

NAD(P)H and O_2 dependant sub-terminal hydroxylation of fatty acids was also detected in cell-free extracts of the fungus *Fusarium oxysporum*, (Shoun *et al*, 1984) and the enzyme involved was later confirmed to be a protein with fused P450 and reductase domains: a eukaryotic counterpart of CYP102A1 (Nakayama *et al*, 1996). This enzyme, termed P450foxy (CYP505), was indeed found to be structurally and functionally identical to CYP102A1, except for three

key differences. Firstly the wild-type P450foxy is membrane-bound, as opposed to the soluble P450 BM-3. Secondly the P450foxy activity was induced by the presence of fatty acid substrates, unlike the P450 BM-3, which in fact seems to be indirectly induced by non-substrate barbiturates. Thirdly, P450foxy seems to favor shorter chain length fatty acids than P450 BM-3 (Kitazume *et al*, 2002a, b; Nakayama *et al*, 1996; Nahri and Fulco, 1987).

Expression of this enzyme in *S. cerevisiae* (Kitazume *et al*, 2000) and *E. coli* (Kitazume *et al*, 2002) resulted in recombinant proteins that were recovered from the soluble fraction, in contrast to its membrane-bound wild-type nature. This was likely due to inefficient post-translational modifications performed by the hosts used.

More eukaryotic self-sufficient P450s have been found, notably P450 Fum6p from *Giberella monoliformis* involved in the biosynthesis of fumonisin, a mycotoxin. This enzyme is the closest counterpart to P450foxy, although the substrates it acts on are probably structurally very different from the fatty acids recognized by P450foxy and P450 BM-3 (Kitazume *et al*, 2002b).

The use of such self-sufficient P450s, and mutants derived from them, is an exciting branch of cytochrome P450 research, and is highly promising in the development of industrial applications using this type of enzyme. Heterologous expression of enzymes in hosts more suited for studying them is a valuable tool in gaining more knowledge and understanding about the proteins, especially new proteins. Expression of human P450s in yeasts such as *Saccharomyces cerevisiae* (Cheng *et al*, 2006), *Pichia pastoris* (Kolar *et al*, 2007), and *Yarrowia lipolytica* (Nthangeni *et al*, 2004); has been successful and has reduced the need for isolation from human tissue. It has led to improvement in the understanding of the substrate specificity and catalytic abilities of the enzymes. Further possibilities, such as enzyme inhibition can also be studied in these systems. The yeasts *Pichia pastoris* and *Yarrowia lipolytica* have been extensively

developed as alternative eukaryotic hosts for heterologous production of intracellular and extracellular proteins. The *Pichia pastoris* system has been highly successful for various proteins, and is commercially available from Invitrogen, USA. In the cases of some proteins however, problems with this system were encountered, including hyperglycosylation and low secretion levels (Cereghino *et al*, 2002; Cereghino and Cregg, 2000).

Thus even though a number of expression systems have been used for the expression of especially eukaryotic P450s, even the most promising ones still have room for improvement. *Y. lipolytica* has proven to be one of the more promising yeast alternative hosts for high level production of active heterologous cytochrome P450s. The *Yarrowia lipolytica* system is also commercially available from Yeastern, Taiwan (http://www.yeastern.com).

1.3. *Yarrowia lipolytica* as a host with high potential for heterologous expression of cytochrome P450 monooxygenases

Briefly introducing *Yarrowia lipolytica*

Yarrowia lipolytica is a non-pathogenic, dimorphic, ascomycetous yeast which is capable of utilizing various hydrophobic substrates as carbon sources. In fact, it is routinely isolated from lipid rich environments including oil fields and food products, such as cheese. It is an obligate aerobe and is one of the best studied 'non-conventional' yeasts to date. A lot of information about the physiology and genetics of this organism has been accumulated (Barth and Gaillardin, 1996; Casarégola *et al*, 1997). This combined with the fact that *Y. lipolytica* has the ability to efficiently utilize hydrophobic substrates makes it, at least at first glance, an attractive host for CYP450 expression (Barth and Gaillardin, 1996; Fickers *et al*, 2004).

Interest in *Yarrowia lipolytica*

Initial interest in this organism was due to its ability to use alkanes as a carbon and energy source, for the production of single cell protein (SCP), as well as organic acids such as citric acid. Its use in several industrial processes has been granted Generally Regarded As Safe (GRAS) status by the American Food and Drug Administration (FDA). Indeed, the non-pathological nature of this organism highlights its potential as a 'green chemistry' catalyst. While these applications for this yeast have not been abandoned, the interest in it has expanded considerably. The applications now reach further into the food industry, especially with regard to the production of flavor and fragrance compounds; and also into bioremediation and fine chemical synthesis, to mention a few (Fickers *et al*, 2005). Furthermore *Y. lipolytica* is well known for its ability to secrete high molecular weight proteins in large amounts, especially proteases and lipases, therefore earning interest as a potential host for heterologous protein expression, especially where protein secretion is desirable (Nicaud *et al*, 2002; Juretzek *et al*, 2001; Pignede *et al*, 2000).

Genetic information and tools available for manipulating *Y. lipolytica*

The complete genome of *Y. lipolytica* strain CLIB122 (E150 of the French *Y. lipolytica* in-breeding lineage), which is split into six chromosomes, has been sequenced by the Génolevures Consortium, (http://cbi.labri.fr/Genolevures/elt/YALI). The complete genomic sequence is also available as a BLAST template on the website of the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST/).

A significant amount of research has been dedicated to optimizing the genetic tools for heterologous gene expression in this yeast. The key to an efficient expression system is a harmonious coalition of expression vectors, host strains

and mechanisms of vector maintenance within host strains, involved in the system.

Selection of a strain can be based on performance characteristics such as growth or protein secretion. The strain should also be devoid of alkaline extracellular proteases if proteins are to be secreted. Lastly and most importantly, it should also be deficient for producing some important nutrient, to serve as an auxotrophic selection marker. Strains can also be customized further for specific purposes. For example in the work done by Picataggio *et al* on *Candida tropicalis*, as well as in our research on *Yarrowia lipolytica*, mutants were created with some of the genes encoding fatty acyl-coA oxidases, important in the *β*-oxidation pathway, deleted. This was done to facilitate the accumulation of intact dicarboxylic acids from alkanes, mono-alcohols or diols (Picataggio *et al*, 1992; Smit *et al*, 2005).

In the case of *Y. lipolytica*, essential components for expression vectors are (i) the expression cassette, (ii) a selection marker, and (iii) elements ensuring maintenance within the host cell; all of which will be briefly discussed in the following sections. Aside from these, the vector should contain a versatile multiple cloning site with sufficient unique restriction site options for gene insertion.

In the case of *Y. lipolytica*, the vectors incorporated are generally auto-cloning shuttle expression vectors, consisting of bacterial and yeast sections. The bacterial portion usually contains only a bacterial origin of replication and an antibiotic resistance gene for selection of bacterial transformants. The vectors are referred to as auto-cloning when the bacterial portion can be separated from the yeast section prior to introduction into the yeast, usually via restriction enzyme digestion. The yeast portion can be isolated by gel electrophoresis and can be used for integration into the host genome (Pignede *et al*, 2000; Nicaud *et al*, 2002). If a transformed strain is destined for industrial application the

presence of a bacterial region is undesirable, since an organism containing bacterial DNA is considered to be genetically modified (Pignede *et al*, 2000).

(i) Expression cassettes

The expression cassette consists of the gene to be expressed (or over expressed,) sandwiched between a promoter and a terminator region. Promoters generally used for high level expression of cloned genes are described in table 1.1.

Table 1.1: Promoters generally used for heterologous expression (adapted from Madzak *et al*, 2004)

Promoter	**Source**	**Inducer(s)**
Inducible promoters		
LEU2	β-isopropylmalate dehydrogenase	Leucine precursor
XPR2	Alkaline extracellular protease	Peptones (pH value)
*POT*1	3-oxo-acyl-CoA thiolase	Fatty acids and their derivatives; alkanes
POX2	Acyl-CoA oxidase 2	Fatty acids and their derivatives; alkanes
POX1 and p*POX5*	Acyl-CoA oxidases 1 and 5	Weakly inducible by alkanes
*ICL*1	Isocitrate lyase	Fatty acids and their derivatives; alkanes; ethanol and acetate
G3P	Glycerol-3-phosphate dehydrogenase	Glycerol
MTP	Bidirectional: metalothioneins 1 and 2	Metallic salts

hp4d	Hybrid promoter, derived from p*XPR2*	Growth phase dependant
Constitutive promoters		
TEF	Translation elongation factor-1α	
RPS7	Ribosomal protein S7	

For some applications, such as when differentiation between cloned and native activities is necessary, inducible promoters are more attractive than constitutive ones. The p*XPR2* promoter is traditionally the most extensively used and studied inducible promoter in the case of *Y. lipolytica*. However, its complicated regulation has limited its practical application. This promoter only functions at a pH above 6 and requires a rich media with high peptone content. It is therefore ineffective in minimal media. Work has hence been dedicated to finding alternative promoters to the conventional p*XPR2* promoter (Juretzek *et al*, 2000).

A comparison was made by Juretzek and co-workers, between the *G3P, ICL1, POT1,* and *POX1, 2* and *5* promoters for their activities and regulatory capacity during growth on various carbon sources. The *lacZ* gene from *E. coli* was used as a reporter gene in single copy transformations of *Y. lipolytica.* The resultant β-galactosidase activities showed that p*ICL*, p*POT1* and p*POX2* were the strongest inducible promoters for expression in this yeast. Furthermore, they found that glucose and glycerol completely repressed the p*POT1* and p*POX2* promoters, and partially repressed the p*ICL* promoter to a basal level. All three of these promoters were strongly induced by oleic acid, and also induced by alkanes, while p*ICL* was also strongly induced by ethanol (Juretzek *et al*, 2000).

In a different study, a hybrid promoter was constructed by combining tandem copies of the UAS1B element from p*XPR2* which is not readily affected by environmental conditions, with a minimal *LEU2* promoter, reduced to only its TATA box. This hybrid promoter was termed hp4d, for 'hybrid promoter' with four

copies of the UAS1B element in the direct orientation. This promoter has proven to be very effective in driving expression, and also quite independent of environmental conditions. It is not affected by carbon or nitrogen sources, or extracellular pH, and does not require peptone in the culture media. It is not truly constitutive as a promoter though, since it drives gene expression which tends to be somewhat growth-phase dependant, starting production at the commencement of stationary phase (Madzak *et al*, 2000; Nicaud *et al*, 2002).

Our research group has focused mainly on the use of the p*POX2* promoter of the peroxisomal acyl-CoA oxidase *POX2* gene and p*ICL* promoter of the anapleorotic isocitrate lyase *ICL* gene. They were both used to drive the expression of a P450 reductase CPR for enhancement of P450 activity, and approximately equal enhancement was observed for both under induction with olive oil (Nthangeni *et al*, 2004; Shiningavamwe *et al*, 2006).

There is information that fragments of the *XPR*2t, *LIP*2t, and *PHO*5t genes have been used successfully as terminator signals (Juretzek *et* al, 2001; Madzak *et al*, 2004; Nicaud *et al*, 2002), but it appears that less consideration goes into the selection of the terminator region than in ideal promoter selection.

(ii) Selection markers

Antibiotic resistance genes are the common choice as selection markers in bacteria. *Y. lipolytica* has however been proven to be resistant to most of the commonly used antibiotics. There have however been successful applications of selection using antibiotic resistance against the few antibiotics which *Y. lipolytica* is sensitive to (e.g. hygromycin, Cordero Otereo 1996), but spontaneous resistance developed after some time, hindering this method (Barth and Gaillardin, 1996).

Thus the use of nutritional requirements is a more practical approach when using *Y. lipolytica* (Madzak *et al*, 2004). The strain to be used should be auxotrophic towards at least one nutrient, a characteristic which can be corrected by complementation using an appropriate vector containing a functional gene, which will restore the prototrophic nature of the strain. Thus only transformed strains will be able to grow on media lacking the particular nutrient.

Many mutant strains of *Y. lipolytica* possessing auxotrophy to different nutrients are available. Heterologous *SUC2* expression from *S. cerevisiael,* as a dominant marker, has allowed transformant selection, due to the inability of wild-type *Y. lipolytica* strains to grow on sucrose as sole carbon source. The marker genes used most frequently though are *LEU2* involved in leucine synthesis and *URA3* involved in uracil synthesis, used to complement strains which have these genes disrupted. A defective allele of the *URA3* marker, *ura3d4*, with its promoter truncated by 11 deletions, was created (Le Dall *et al*, 1994). It was designed to allow multiple integrations into the genome at specific target sites (discussed in the following section) since multiple copies of the modified marker are required for restoration of the prototrophic phenotype. It was in fact observed that copy numbers of 10-12 were required for complete complementation of the disrupted *ura3* gene, with lower copy numbers resulting in proportionately inferior growth of the yeast. Successive cultivations in selective media caused an amplification of the copy numbers, stabilizing around 10-12 copies (Juretzek *et al*, 2001). On the other hand, it was found that high expression levels of some proteins resulted in toxicity to the cells, and multiple integrations were only stable when the expression of the proteins in question was not induced. Deamplification of the copy numbers in such cases occurred, also stabilizing around 10-12 copies (Le Dall *et al*, 1994).

This defective marker has since been successfully used for multiple integrations of homologous and heterologous genes (Juretzek *et al*, 2001; Nicaud *et al*, 2002; Nthangeni *et al*, 2004). This defective allele is very effective for obtaining multiple

copy integration using replicative vectors, or integrative vectors for different genomic target sites, and even for the single copy integrative *XPR2* locus. One negative aspect using the defective marker is a dramatic decline in transformation efficiency compared to single copy integrants (Juretzek *et al*, 2001; Le Dall *et al*, 1994).

(iii) Integration targets

Shuttle vectors can be divided into two classes based on their mechanism of maintenance in host cells, namely episomal vectors and integrative vectors.
Episomal plasmids do not naturally occur in *Y. lipolytica*, but replicative plasmids using chromosomal origins of replication can be designed. These are based on several autonomously replicating sequences (ARS) isolated in *Y. lipolytica*. The plasmids based on these ARS elements are however impractical for higher amplification of gene expression, since the copy numbers of genes expressed using these vectors are limited to 1-3 copies per cell. Another drawback is the need for maintenance by selective pressure (Madzak *et al*, 2004).

Integrative plasmids facilitate integration by homologous recombination. This can occur at sites in the genome such as *LEU2*, *URA3* and *XPR2*, by linearizing the homology region in the plasmid (Barth and Gaillardin, 1996). Alternatively, some strains have been transformed with a *URA3* integrating vector containing DNA from the pBR322 vector, creating a docking site for direct integration using pBR based expression / secretion vectors (Nicaud *et al*, 2002; Madzak *et al*, 2000).
In over 80% of cases relying on either of these means of homologous recombinations, only a single complete copy of the vector will be integrated into the genome at the desired site (Barth and Gaillardin, 1996).

For achieving multiple integrations into genomic DNA, repetitive genomic elements need to be targeted for homologous integration, with vectors containing defective selection markers (described in previous section). There are two such

targets in *Yarrowia lipolytica*, namely tandemly repeated ribosomal DNA (rDNA) sequences, and the dispersed repeated Ylt1 retrotransposon.

The rDNA consists of an omnipresent gene cluster G (G-Unit) and additional clusters P1 and P2, present in some strains. These clusters were detected on five of six chromosomes. A strain derived directly from the wild-type showed a similar rDNA pattern to a zeta containing inbred French laboratory strain, with G units exceeding P units. Targeting such a major rDNA repeat such as the G-Unit, resulted in transformants containing up to 60 copies of the plasmid, although in these cases, the transformants became unstable when protein expression was induced, leading to deamplification, until copy number stabilized around 10-13 copies (Le Dall *et al*, 1994; Juretzek *et al*, 2001).

The Ylt1 retrotransposon is bound by a long terminal repeat (LTR) named the zeta element, a highly conserved region which also exists as a solo element. The numbers of Ylt1 and solo zeta elements vary per strain, but at least 35 copies of Ylt1 and in excess of 30 solo zeta elements (up to 50-60 copies) have been observed for a single haploid genome, resulting in 65-95 potential target sites for integration (Barth and Gaillardin, 1996; Juretzek *et al*, 2001). When zeta elements were targeted for multi-copy integration using vectors containing the defective *ura3d4* marker, the copy numbers again stabilized around 10-12 copies, the number required for complete complementation of the uracil auxotrophy. As was mentioned earlier, the defective marker proved to be the copy number determinant, regardless of the integration target used.

It is important to note that these zeta elements are not present in all *Y. lipolytica* strains, and are absent in most wild-type strains. Only strains from inbreeding lines that have at some point involved crossing with the American (Wickerham) strain YB423, from which these elements seem to have originated, contain zeta elements (Barth and Gaillardin, 1996; Juretzek *et al*, 2001).

An interesting observation was that using a plasmid with zeta elements for integration in a strain devoid of zeta elements still gave integration. Such non-homologous integration is far more dispersed, which seems to increase copy number stability. The transformation efficiencies of the heterologous integrants tend to be much lower than with homologous integrations though. The other problem is that since the integration is essentially random, the integration sites are unpredictable, essentially creating mutations with unpredictable genotypic outcomes. The cells may suffer adverse consequences from some of these integrations, should the integrations occur at inappropriate sites. Nonetheless, multiple copies were integrated into the genome of zeta free strains using vectors equipped for zeta integration and containing the defective ura marker allele (Pignede *et al*, 2000; Juretzek *et al*, 2001; Nicaud *et al*, 2002; Nthangeni *et al*, 2004).

This random, non-homologous integration strategy has been used for the generation of tagged mutants in *Y. lipolytica*, as a type of directed evolution. These mutants were affected in hydrophobic substrate degradation for the study of genes involved in these pathways. A mutagenesis cassette (MTC) was constructed, containing the non-defective *ura3d1* allele as a selection marker, flanked on either side by zeta elements. The MTC was used to transform a zeta free strain of *Y. lipolytica*. Transformants were initially screened for growth in uracil free media, after which successful isolates were grown on minimal media containing hydrophobic substrates as sole carbon source. Disrupted genes were identified using convergent and divergent PCR techniques (Mauersberger *et al*, 2001; Thevenieau *et al*, 2007).

In summary, multiple copy transformants of comparable copy numbers were obtained whether rDNA, zeta elements, or the single copy XPR2 locus were targeted, indicating that the defective marker *ura3d4* is the determinant factor of copy number, rather than the integration site targeted. Tandem repeats were obtained when the homologous recombination targets XPR2 locus and rDNA

were used, while homologous integration to the zeta elements showed predominantly tandem but also occasional dispersed integration. Non-homologous integration using the zeta element based vectors gave integration which was far more random and dispersed. The stability of the gene copy number in the transformant was apparently better in the cases where integration was more dispersed than in tandem. An example of a typical vector used in our group, based on zeta element integration is given in figure 1.5.

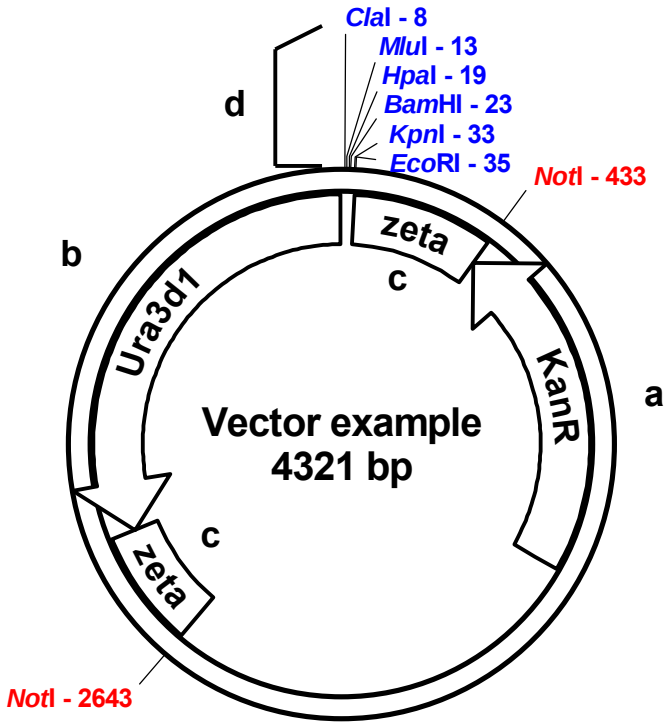


Figure 1.5: Illustration of a typical shuttle vector used for genomic DNA integration into *Y. lipolytica* based on zeta element integration.

The selection markers in this vector are kanamycin resistance for bacterial sub-cloning (a), and the uracil marker for selection of yeast clones (b). The example contains the non-defective *ura3d1* marker, which allows single copy integration, but can be replaced by the defective allele *ura3d4* to obtain multiple copy integrants. The bacterial fragment can be separated from the yeast component using the *Not*I restriction sites which can be seen flanking the zeta elements (c).

These zeta elements offer a mode for integration into *Y. lipolytica* host strains, either homologously or heterologously, depending on the host strain used. Finally, the multiple cloning site (d) contains a variety of restriction sites unique to the plasmid, allowing insertion of genes of interest, such as P450 genes and promoters to control their expression.

Nicaud *et al* (2002) successfully scaled up protein production by *Y.lipolytica* using multi-copy integration. For a homologous extracellular lipase LIP2p, they doubled expressed protein activity in shake flasks obtained with single copy integration (via pBR docking) by using multiple copy integration (via non-homologous recombination) achieved with the defective *ura3d4* allele. Furthermore, they raised the multiple copy protein activity moving from shake flasks to batch cultivation by almost 6-fold, and a further nearly 8-fold increase in the shift from batch to fed-batch cultivation. Hence the fed-batch cultivation using a defective marker for multiple integrations improved the protein activity over single copy shake flask cultivation by nearly 100 times (91-fold). For heterologous production of leucine amino peptidase II from *Aspergillus oryzae* in *Yarowia lipolytica*, there was an 8-fold increase in batch cultures shifting from mono-copy to multi-copy integration. There was a further 11-fold activity increase for multi-copy integrants grown in fed-batch compared to batch cultivation. Hence, in both cases, production of homologous and heterologous proteins, the activity was increased by almost 100 times when gene amplifications using a defective selection marker were combined with fed-batch cultivation, even under conditions described as sub-optimal (Nicaud *et al*, 2002).

1.4. Motivation for selecting *Y. lipolytica* as a host for heterologous P450s

Escherichia coli was initially, and is still, used for heterologous expression of CYP450s, but despite bacterial cells having the advantages of producing fairly large amounts of protein in a short time and under inexpensive conditions, they

were shown to have certain drawbacks for expression of eukaryotic P450s. A major limitation of *E. coli* as a host for these purposes is the lack of an endoplasmic reticulum, in which the eukaryotic P450s reside and receive their reducing equivalents. Development of CYP450-CPR bicistronic or biplasmid expression systems resulted in loss of cell viability and reduction in catalytic activities of the expressed proteins. Bacteria are further limited in their inability to perform the post-translational modifications required by the eukaryotic proteins. High level P450 expression could only occur once further modifications of 20-30 N-terminal amino acid residues were done (Sakaki and Inouye, 2000; Iwata *et al*, 1998).

Yeasts are attractive alternatives as hosts for such expressions, since they combine the complex machinery for post translational modification (e.g. glycosylation) of more complex eukaryotic protein systems, with the simplicity of growing and manipulating unicellular organisms. Therefore optimization of the systems involved is required, and this has followed two basic routes: improving the initial, moderately successful, *Saccharomyces cerevisiae* system, or finding alternative hosts to perform these expressions. Müller *et al* (1998) investigated alternative hosts to *Saccharomyces cerevisiae* for heterologous protein expression. They compared the capacity of *S. cerevisiae* and four other yeasts to express and secrete six fungal enzymes in their active forms. They found *Yarrowia lipolytica* to be the most efficient host for these purposes. It also had the most reliable performance reproducibility. The other yeasts tested were *Hansenula polymorpha*, *Klyveromyces lactis* and *Schizosaccharomyces pombe* (Müller *et al*, 1998).

Indeed *Y. lipolytica* has been successfully used for the heterologous expression of a number of proteins of eukaryotic origin. There are considerable differences in glycosylation patterns required amongst eukaryotic proteins, for example, between yeasts and mammals (Cereghino and Cregg, 2000). Aside from merely performing post-translational modifications, *Y. lipolytica* was shown to adhere

closely to the natural pattern of glycosylation of the specific protein, even for mammalian proteins. A good example is the case of prochymosin, which has two potential sites for glycosylation, but is not glycosylated by its natural hosts, nor by *Y. lipolytica* as a host for this protein (Madzak *et al*, 2000; Madzak *et al*, 2005; Swennen *et al*, 2002; Madzak *et al*, 2004).

Complementary to these factors, provisional optimization conditions for production of homologous and heterologous proteins in fed-batch as well as batch cultures have been investigated, with regards to media, culture conditions and genetic factors employed (Nicaud *et al*, 2002).

Hydrophobic substrate metabolism by *Y. lipolytica*

An additional reason for considering *Y. lipolytica* as a host for CYP450 expression is the fact that this yeast is very efficient at metabolizing hydrocarbons. This happens through a series of steps in various subcellular compartments within the cell, by various enzymes belonging to different multigene families. These include lipases / esterases (*LIP* genes), cytochrome P450s (*ALK* genes), and acyl-CoA oxidases (*POX* genes). Aside from the various degradation enzymes involved, there are also transport proteins which play an important role in these processes. When considering the hydroxylation of hydrophobic substrates by whole cells expressing CYP450s, it is important to keep the endogenous mechanism of hydrophobic substrate metabolism in mind.

In the case of alkane metabolism, the process starts with substrate uptake into the cells either through direct cell wall adherence or by emulsification by surfactants (Fickers *et al*, 2005). The initial interactions presumably occur at protrusion complexes which have been observed on the cell surface of alkane grown *Yarrowia lipolytica* and *Candida maltosa* cells (Thevenieau *et al*, 2006; Mlickova *et al*, 2004). Specific transporters have been identified which have been shown to be involved in substrate uptake, and are thought to form part of these

protuberances. These transporters are encoded by genes *ABC1* for longer chain alkanes, e.g. C14 and C16, and *ABC2-4* for shorter chain length alkanes, e.g. C10 and C12. It has not yet been confirmed if all of the genes *ABC2-4* are involved in shorter chain alkane uptake, or whether they also play a role in the exportation of alkanes as a regulation mechanism for intracellular alkane concentration. These transporters are presumably also involved in the channeling of the alkanes into the endoplasmic reticulum (ER) (Thevenieau *et al*, 2006).

Once in the ER of the cell, the alkanes are terminally hydroxylated by cytochrome P450 monooxygenases to fatty alcohols. Although 12 alkane hydroxylating P450s have been identified in *Y. lipolytica* (CYP52F1-CYP52F12), only two have been shown to be essential for alkane conversion to their respective fatty alcohols. These are encoded by *YlALK1* (CYP52F1) for assimilation of decane (C10) and longer molecules, and *YlALK2* (CYP52F2) for assimilation of molecules longer than dodecane (C12) (Iida *et al*, 2000).

The produced alcohols are then converted to their corresponding fatty aldehydes. This can be achieved by one of two routes: by an $NAD(P)^+$ dependant alcohol dehydrogenase in the ER (usually longer chain alkanes), or by a peroxisomal, hydrogen peroxide forming fatty alcohol oxidase, after transportation into the peroxisomes (usually shorter chain alkanes). The produced aldehydes are converted to fatty acids by $NAD(P)^+$ dependant fatty aldehyde dehydrogenases present in both the ER and the peroxisomes (Thevenieau *et al*, 2006).

Surprisingly, although fatty alcohol oxidase activity has been observed in the wild-type *Y. lipolytica* strain H222 (Mauersberger *et al*, 1992,) no gene encoding for fatty alcohol oxidase was found in a search of the genome of the sequenced *Y. lipolytica* strain E150 (CLIB122) (Smit, personal communication). At least one putative fatty alcohol dehydrogenase (FADH) gene and four putative fatty

aldehyde dehydrogenase (FALDH) genes have been identified in the *Y. lipolytica* genome (Fickers *et al*, 2005).

An important observation to keep in mind regarding the role of fatty alcohol oxidases and fatty aldehyde dehydrogenases in fatty alcohol oxidation is the complete conversion of *n*-alkanes to α, ω-dioic acids by only the heterologous CYP52A3 from *C. maltosa* expressed in *S. cerevisiae*. This single enzyme catalysed conversion was shown to occur via all of the intermediates, (i.e. the mono-alcohol, the aldehyde, the mono-acid, the hydroxy acid, and the diol) (Scheller *et al*, 1998). While the ω-hydroxylation of fatty acids has been observed for some of CYP52 enzymes from *Y. lipolytica* (encoded by genes *ALK3*, *ALK5* and *ALK7*) (Hanley *et al*, 2003), it is yet to be confirmed whether or not these *Y. lipolytica* P450s can emulate cascades similar to those observed for the CYP52A3 from *C. maltosa*. Using whole cells from strains of both *Candida tropicalis* (Picataggio *et al*, 1992) and *Yarrowia lipolytica* (Smit *et* al, 2005) with disrupted β-oxidation pathways, accumulation of dioic acids was observed from alkanes and other intermediates. The use of whole cells, however, made the full extent of CYP52 involvement uncertain, due to the presence of other enzymes involved in intermediate conversions.

The longer chain fatty acids produced in the ER are activated by the cytoplasmic ATP-dependant acyl-CoA synthetase ACSI, which has long chain specificity. The corresponding Acyl-CoA esters are transported into the peroxisomes via an ABC transporter encoded by *PXA1 / PXA2*, followed by degradation via proteins encoded by the genes *POX1-6* of the cells peroxisomal β-oxidation pathway. The shorter chain molecules, which are generally converted to their corresponding fatty acids within the peroxisomes, are activated to their respective Acyl-CoA esters by the ATP-dependant acyl-CoA synthetase ACSII, which receives its peroxisomal ATP via an adenine nucleotide transporter Antlp. Longer chain fatty alcohols which were directly transported into the peroxisomes are also activated (for uptake in the β-oxidation pathway,) by ACSII, which has broader substrate

specificity than ACSI. Short and medium chain fatty acids produced in the ER follow the same route after transportation into the peroxisomes via an as yet unidentified transporter (Thevenieau *et al*, 2006; Fickers *et al*, 2005; Nicaud *et al*, 1998). A scheme illustrating these routes and principles is given in figure 1.6.

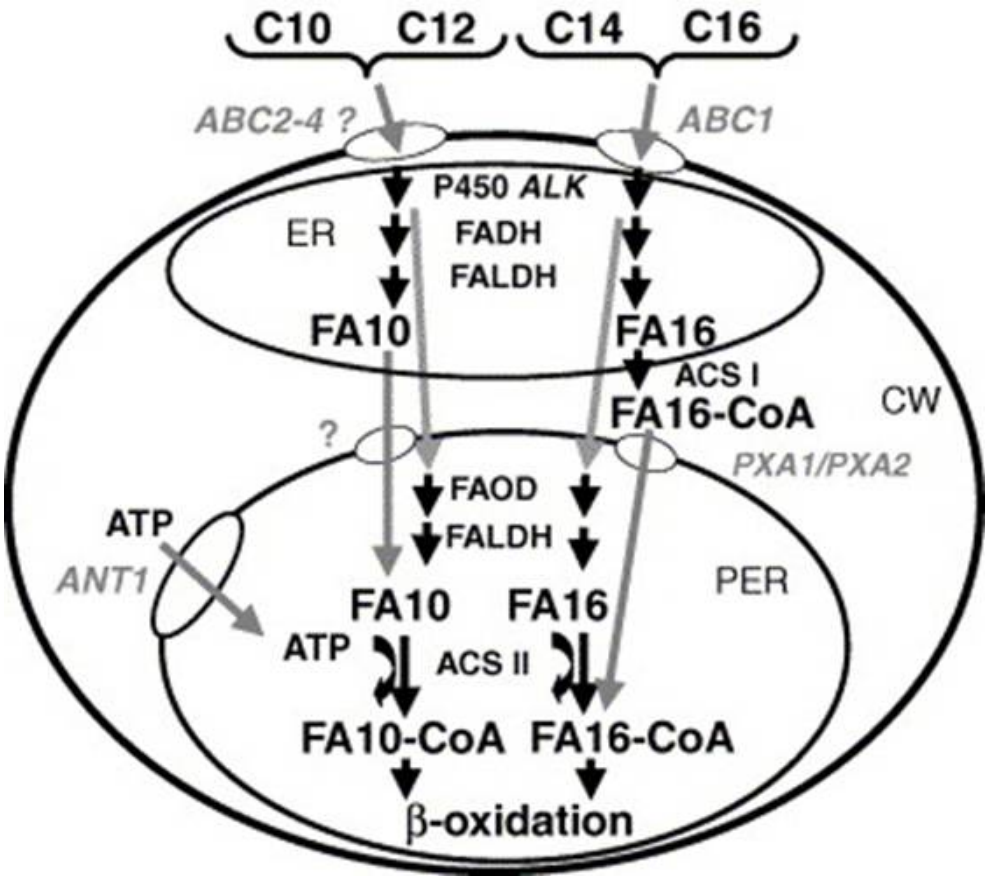


Figure 1.6: Schematic representation of the proposed alkane uptake, transport and degradation routes in *Y. lipolytica*. The various proteins involved in transport and degradation of substrates and intermediates discussed in the text are illustrated in perspective (from Thevenieau *et al*, 2006).

The factors discussed thus far highlight *Y. lipolytica* as a strong candidate for heterologous expression of P450 enzymes in high quantities. It has at least 12 P450s of its own and thus already contains all the required sub-cellular

organization, endoplasmic reticular proliferation and efficient electron transport systems required by these enzymes. It also has all the necessary transport systems required for the uptake and intracellular channeling of hydrophobic substrates (Thevenieau *et* al, 2006; Fickers *et al*, 2005).

1.5. Examples of P450 expression in *Y. lipolytica*

At least 50 heterologous proteins have been successfully expressed in *Yarrowia lipolytica*. The number had increased in the period between 1991 and 2004 from 5 to 42 (detailed list available in Madzak *et al*, 2004). The proteins are diverse in function, including enzymes (which in turn include cytochrome containing enzymes), antibodies, metalloproteins and classical reporter proteins. They originate from viruses, eubacteria, fungi / yeasts, plants and mammals, thus illustrating the great diversity of proteins that this yeast is capable of expressing correctly. Most of these proteins could be produced at a level of milligrams per liter in shake flasks, without the 10-20 fold increase that can be introduced using multi-copy integration (Madzak *et al*, 2004).

At least five heterologous cytochrome P450 monooxygenases have been expressed in *Y. lipolytica*, as well as two homologous P450 enzymes. Of the five heterologous P450 proteins, two are yeast proteins, one is a plant enzyme, and two are mammalian proteins.

Eukaryotic P450s are generally membrane bound (usually of the endoplasmic reticulum) and are hence not secreted. Activities can therefore be determined using cell free extracts (specifically microsomal fractions,) or whole cells. Eukaryotic P450s belonging to classII are dependant on a specific NAD(P)H: cytochrome P450 reductase (CPR) which directly reduces the P450 active site (Werck-Reichart and Feyereinsen, 2000). This multi-protein component system, together with the requirement for continuous supply of reducing equivalents,

complicates the use of isolated enzymes (or microsomal fractions) when studying this class of P450. Although the problems are partially alleviated when isolated self-sufficient P450s are used, such as P450 BM-3 (CYP102), regeneration of expensive NAD(P)H is still required for *in vitro* applications. Stoichiometric supply of NAD(P)H for large scale industrial applications is economically impractical. Further problems are often experienced with solubility when using membrane bound P450s, and removing them from their membrane environment tends to result in loss of activity.

These problems can be avoided by allowing the reactions to occur *in vivo*, using whole cells. Whole cells are used as either resting or growing cells. Resting cells are obtained by harvesting already induced cells, resuspending them in buffer and then adding substrates for bioconversion. Using growing cells, substrates are added to the cultivation medium in which cells are still active.

Bourel *et al* (2004) cloned a fatty acid hydroperoxide lyase from green bell pepper into *Y. lipolytica* under the control of the *POX2* promoter, using non-homologous zeta integration into the genome of *Y. lipolytica* strain P01d. Contrary to the general mechanism of P450s, this enzyme uses hydrogen peroxide as a source of activated oxygen instead of molecular oxygen, and is not dependant on reducing equivalents for its activity (figure 1.7.).

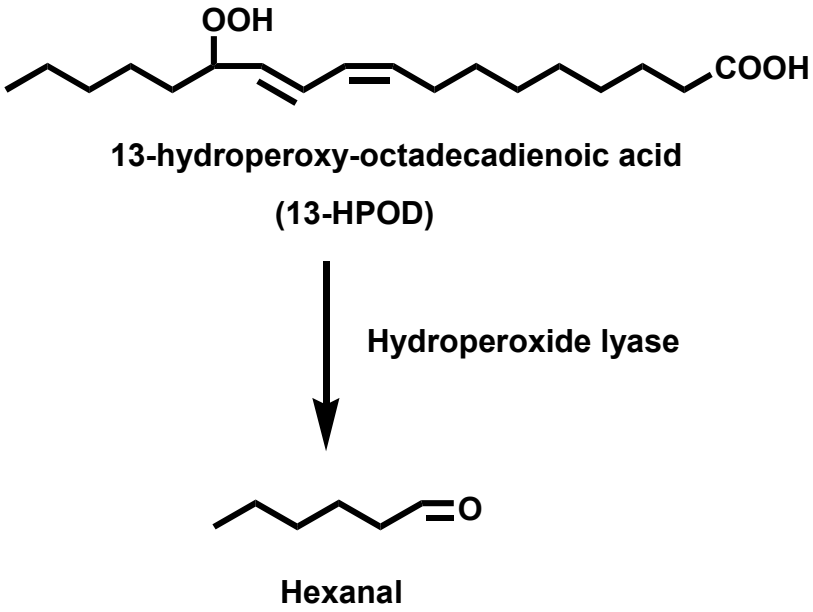


Figure 1.7: Formation of hexanal from a hydroperoxide derivative of linoleic acid as an example of hydroperoxide lyase activity.

Activities in cell free extracts were compared to whole cell activities, for resting or growing cells. Although the expected reaction was obtained in each case, the lowest quantity of product was obtained using resting cells (220 mg/L), while the growing whole cells produced a higher concentration of product (350 mg/L) than the cell-free extract (312 mg/L), although over a longer reaction time. Potential routes for optimization of this heterologous expression would be expressing the enzymes using vectors allowing multi-copy integration, and by disrupting fatty aldehyde dehydrogenases in the cell (Bourel *et al*, 2004).

Human cytochrome P450 CYP1A1 was cloned into *Y. lipolytica* under the control of the *POX2* promoter in mono-and multi-copy, using non-homologous zeta integration into the genome of *Y. lipolytica* strain P01d. Activity for hydroxyresorufin conversion to resorufin was tested after induction using olive oil, and little activity of the cloned enzyme was observed using resting cells, even in multi-copy cases (Nthangeni *et al*, 2004).

Additional copies of the homologous CPR gene were cloned into some of these strains. Strains co-expressing both hetrologous enzymes had dramatically increased CYP1A1 activity, compared to strains expressing only the hetrologous P450. This held even though only one or two additional CPR copies were cloned, compared to multiple CYP copies. As expected, no CYP1A1 activity was observed when the additional CPR was expressed without the co-expression of CYP1A1. The increase in activity was approximately the same whether the CPR gene was cloned under the *POX2* promoter or the *ICL* promoter using olive oil for induction. This indicated that the CPR gene itself is the major determinant of increased P450 activity (Nthangeni *et al*, 2004). The results also indicated that little difference in strength between these promoters, which is in agreement with reported results showing approximately equal heterologous β-galactosidase activity in *Y. lipolytica* whether the LacZ gene was cloned under the *POX2* or *ICL* promoter (Juretzek *et al*, 2000).

A possible explanation could be that the competition between the cloned cytochrome P450s and the cell's native hemoproteins for reducing equivalents, leading to an overall decrease in the availability of the reducing equivalents. The CPR could be incapable of keeping up with the increased demand for these reducing equivalents, preventing the cloned P450s from reaching optimal activity.

During investigation of dioic acid production from alkanes in β-oxidation disrupted *Candida tropicalis* strains, an increase in production levels was also observed in strains when there was co-expression of additional CPR reductase and P450Alk1 genes (Picataggio *et al*, 1992).

The results also complemented other research in which the effect of the amount of yeast reductase proteins co-expressed with mammalian P450s in *S. cerevisiae* was tested (Truan *et al*, 1993; Murakami *et al*, 1990). The highest heterologous CYP1A1 activity by *Y. lipolytica* was 50-fold higher than the highest activity

obtained by heterologous expression of CYP1A1 in *S. cerevisiae*, (Nthangeni *et al*, 2004).

Another mammalian P450, bovine cytochrome P45017α, was functionally expressed in *Y. lipolytica* under the control of the *ICL* promoter, as well as in *S. cerevisiae* under the control of the GAL10 promoter, for comparison. The enzyme converts progesterone to 17α-hydroxyprogesterone. The defective *ura3d4* marker was used for multiple integrations into the genome of *Y. lipolytica* strain P01d, based on both zeta element (non-homologous using P01d) and rDNA integration (homologous) targets, separately. The gene copy number was far higher for *S. cerevisiae* cells than for *Y. lipolytica* cells and the microsomal protein content of the two yeasts, determined using carbon monoxide difference spectrum and western blotting, duly corresponded. Despite this, the *Y. lipolytica* cells induced with either ethanol or hexadecane, had far higher activities than the galactose induced *S. cerevisiae* cells, using microsomal fractions as well as whole cells. Furthermore, although hexadecane-induced microsomal activities were only slightly higher than for ethanol induced microsomes, whole cell activities were far higher during hexadecane induction than during ethanol induction. These results pose interesting questions about the effect of ethanol on other systems in whole cells involved in hydrophobic substrate metabolism (Juretzek *et al*, 2000b).

Expression of four more cytochrome P450s in *Y. lipolytica* has been attempted in our research group. Two of these were homologous alkane hydroxylases CYP52F1 and CYP52F2 (Setati, 2004; unpublished data). Shiningavamwe (2004) cloned CYP53B1 from *Rhodotorula minuta* and CYP557A1 from *Rhodotorula retinophila* into *Y. lipolytica.* The first gene encodes a benzoate-*para*-hydroxylase exhibiting activity completely absent in *Y. lipolytica* (figure 1.8.); while the latter encodes a putative alkane and fatty acid hydroxylase, expected to exhibit activity similar to the activity of CYP52F1-12. Both enzymes were cloned

under the *POX2* promoter in single and multi-copy, using homologous zeta integration into the genome of *Y. lipolytica* strain E150 (Shiningavamwe, 2004).

The benzoate-*para*-hydroxylase from *R. minuta* serves as an excellent reporter gene for P450 expression in *Y. lipolytica*, for the following reasons:

(i) benzoate-*para*-hydroxylase activity has not yet been reported for wild-type *Y. lipolytica*, and no significantly similar activity is found in this yeast;
(ii) the product, *p*-hydroxybenzoic acid (pHBA), is highly UV-active, hence easily visualized by rapid thin layer chromatography (TLC) analysis;
(iii) pHBA is not further metabolized by this yeast;
(iv) benzoic acid (BA) is a fairly inexpensive substrate, hence viable for larger scale trials;
(v) at pH >7 BA and pHBA are non-toxic to *Y. lipolytica* at concentrations up to 40 mM;
(vi) BA and pHBA are reasonably water soluble, allowing easier dispersion in media, improved access to the cells and better sampling

(Shiningavamwe *et al*, 2006).

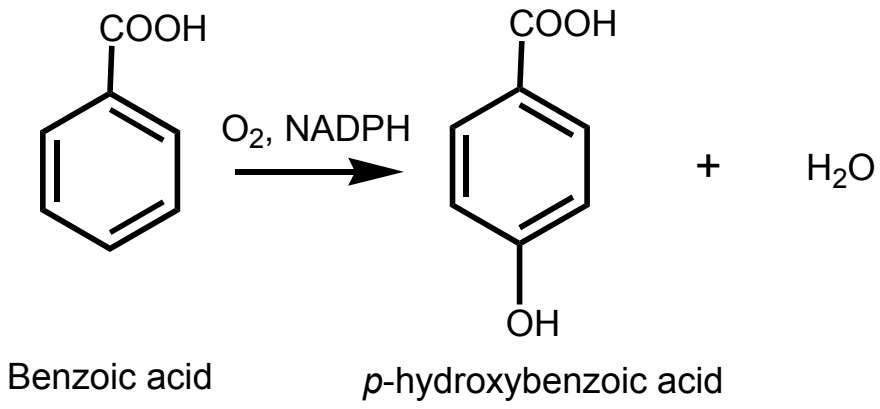


Figure 1.8: Benzoate-para-hydroxylase reaction.

Biotransformation of benzoic acid using growing *Y. lipolytica* whole cells was enhanced when additional CPR genes were cloned into the most promising CYP53 transformants. The CPR genes were cloned in single copy under the *ICL* promoter, using the same method of integration as for the CYP53 genes. Although the activities obtained were still lower than the observed wild-type activity, this is a common observation for heterologous expression (Duetz *et al*, 2001). The CYP53 activities observed were however higher than the hydroxylation activities obtained for the previously mentioned human CYP1A1 heterologously expressed in *Y. lipolytica*, under comparative analytical assay conditions.

A problem encountered when using the *POX2* promoter for P450 expression in *Y. lipolytica* for whole cell biocatalysis, is that at least 4 of the 12 P450s naturally present in this yeast are also induced by fatty acids, which are generally used to induce the *POX2* promoter (Fickers *et al*, 2005; Iida *et al*, 2000). Hence competition arises between the cognate background P450s and the cloned P450s for the CPR and also for reducing equivalents within the cell when the inducer used is also an inducer and substrate of wild-type P450s (Fickers *et al*, 2005).

In the case of CYP53, the activity was low but still sufficient for convenient activity screening since the wild-type has no benzoate-*para*-hydroxylase activity of its own. Cloning of additional reductase genes improved the situation somewhat. The problems are however augmented for an enzyme such as the putative fatty acid and alkane hydroxylase CYP557, since this activity is expected to be similar to the activities of the wild-type P450 enzymes in this yeast. The problems encountered when using such an enzyme are not confined to competition for reducing equivalents, but now extend to a need for differentiation between cloned and wild-type activities. Selective induction of cloned P450s

becomes necessary, and/ or deletion of native P450 genes (Shiningavamwe *et al*, 2006; Obiero, 2006).

Attempted selective induction of the *POX2* promoter in our research group without induction of the background P450s, was unfruitful. The compounds suspected of not being able to induce the wild-type genes, were also unable to induce the *POX2* promoter (Obiero, 2006). Initial experiments using hexylbenzene as a substrate gave some hope that the phenylalkanoic acids formed after hydroxylation of the alkyl chain could selectively induce the *POX2* promoter (figure 1.9.).

Due to the dramatic improvement of P450 activity prompted by additional copies of the CPR, an E150 strain was transformed with a single copy vector containing the CPR gene under the control of the *ICL* promoter (Shiningavamwe, 2004). This strain, TVN496, was then used as a common ancestral strain for further cloning of P450s. Multiple copies of the CYP557A1 as well as the *Y. lipolytica* alkane hydroxylases CYP52F1 and CYP52F2 were then cloned into this strain (Shiningavamwe, 2004; Setati, 2004, unpublished results). Unfortunately, characterization of these strains was limited by the absence of a proper negative control strain. One of the multi-copy CYP53 strains transformed with an additional CPR copy was used as a negative control, since the auxotrophies were remedied in a similar way, and the cloned P450 would be unresponsive to the substrates tested. When ethanol was used to induce the cloned CPR and hexylbenzene was added as a substrate, the strains with cloned alkane hydroxylases consistently gave higher activities of phenylacetic acid production than the makeshift negative control strain TVN91. These results required cautious interpretation though, since the strains expressing the alkane hydroxylases were derived from a common ancestor, while the CYP and additional CPR genes were cloned in a different order in the TVN91 strain. Furthermore, an inducible P450 could be expressed in this strain, whether or not it would have activity for the substrates used, potentially lowering the energy

available for wild-type activity. When phenyldecane and phenyldodecane were used as substrates, no distinction could be made between the different strains. After induction with fatty acids (stearic acid and oleic acid,) and addition of hexylbenzene as substrate, distinction between the activities of the different strains was also impossible (Obiero, 2006; Van Rooyen, 2005).

A derivative of *Y. lipolytica* strain Po1d, a zeta element free strain, was created which had its β-oxidation partially blocked by deletion of four of the six fatty acyl coA oxidases crucial in this pathway. An additional copy of the CPR gene was cloned into this strain under the *POX2* promoter. Multiple copies of CYP557A1 and single copies of CYP52F1 and CYP52F2 were cloned under the *POX2* promoter into the resultant strain. These strains were tested for dioic acid accumulation from dodecane, stearic acid, oleic acid, hexylbenzoic acid and nonyloxybenzoic acid. Improved activity could not be consistently observed in the strains with cloned CYP450s compared to the negative control strains (Van Rooyen, 2005; Obiero, 2006).

1.6. Concluding remarks and aims of this study

From the above information, the vast potential of the *Y. lipolytica* expression system is evident. A diverse range of proteins, including cytochrome P450s, have been successfully expressed in this yeast from their respective hosts, generating proteins of significant structural and functional similarity to their wild-type counterparts. It has been shown to be superior to many other types of yeast tested, in terms of consistent production, secretion and modification of heterologous proteins, while its GRAS status enhances its allure.

Considerable advances have been made in the development of the genetic tools for heterologous gene expression in *Y. lipolytica*. The development of the multiple integration system together with the exploitation of genomic integration

targets have resulted in great progress in increasing heterologous protein expression in modified strains of this yeast. The productions have been further increased by combining such gene amplifications with fed-batch cultivations. Numerous strong promoters have been identified for regulation of heterologous gene expression in this yeast, notably the constitutive *TEF* promoter, and the inducible *POX2*, *ICL*, and the synthesized hybrid promoter hp4d. The creation of auto-cloning vectors in which the bacterial moiety of the plasmid can be removed prior to yeast transformation, makes it even more attractive for industrial applications.

Although the number of heterologously produced proteins by *Y. lipolytica* falls short of production by *E. coli* in many cases, it is better equiped to deal with the expression of more complex eukaryotic proteins. In regard to the lower whole cell activities observed for heterologous P450s in *Y. lipolytica* than in the natural hosts, further optimizations of the current systems need to be investigated. Ideally the yeast should focus much of its energy on heterologous expression, at the expense of neglecting some of its natural proteins. A system which will not induce wild-type proteins, or better still suppress them, would be beneficial for this cause. This is understandably tricky for inducing production of proteins similar to those occurring naturally in the yeast, such as cytochrome P450s. Nevertheless, whole cell biotransformations have generally been, and currently still are, the method of choice in our research group, whether testing natural activity (Smit *et al*, 2005; Van Dyk *et al*, 1998) or activity of heterologously expressed enzymes (Shiningavamwe *et al*, 2006; Obiero 2006).

At our disposal we have a series of vectors containing *POX2* and *ICL* promoters, as well as the single copy *leu* and *ura3d1* markers and the multi-copy *ura3d4* allele of the *URA* marker.

Results obtained by Van Rooyen (2005) and Obiero (2006) indicated that ethanol which is capable of inducing the *ICL* promoter, also represses the wild-type

P450s. Should that be the case, the cell will not waste any of its energy or reducing equivalents on the wild-type P450s, and can focus solely on the P450s cloned under the *ICL* promoter. Therefore the *ICL* promoter appears to be a good alternative to the *POX2* promoter for driving expression of heterologous enzymes during whole cell biocatalysis using *Y. lipolytica*, especially when the cloned genes have activities similar to those of native P450s. Cloning of P450s under the control of the *ICL* promoter for comparison with cloning under the *POX2* promoter thus became the first goal of this study.

Evidence of the importance of the CYP: CPR ratio for optimal activity is clearly demonstrated by the examples discussed. While *Y. lipolytica* produces P450: CPR functional complexes in a 1:1 ratio, there is a number of different P450s produced by this organism, in a 10 to 20 fold excess over the CPR proteins (Backes and Kelley; 2003). The essential P450: CPR ratio is dependant on and is specific for each individual P450 (Truan *et al*, 1993). Hence, although the CPR seems to be indiscriminate of the kind of the P450s it supplies (within P450 family borders), it appears to be continuously occupied with supplying reducing equivalents to all the native P450s. The cloning of foreign CYP genes into *Y. lipolytica* thus necessitates the cloning of additional CPR genes as well to help the system cope with the increased demand for reducing equivalents. Strains in which additional copies of the CPR have been cloned would thus be used for transformation in this study. We have access to *Y. lipolytica* strains derived from E150 (TVN496) as well as PO1d (FT-120), which have additional CPR genes cloned. The FT-120 strain has all of its fatty acyl CoA oxidase genes deleted, rendering its β-oxidation completely blocked. This strain would then be an ideal strain for evaluation of dioic acid accumulation in the presence and absence of cloned P450s.

More suitable negative control strains for efficient comparisons to be made in this study were still outstanding. The strains containing additional CPR copies could be transformed with an ‘empty’ vector, containing no CYP or CPR gene, to

restore the prototrophic phenotype of the strain. The FT-120 strain, which already contained an additional CPR gene under the control of the *POX2* promoter, and the E150 ancestral strain, could be transformed with two 'empty' vectors which would complement both defective markers. Another additional CPR copy could be cloned into the FT-120 strain to test the effect of additional copies, since the marker was available.

The efficient reporter gene of CYP53 would be used to evaluate the modified expression system. From there the system would be used for examination of the putative alkane and fatty acid hydroxylase CYP557, for which selective induction is not negotiable when using *Y. lipolytica* as a host. If sufficient selective induction could occur, crucial information about the activity of this enzyme could be gathered.

The FT-120 strains expressing the putative fatty acid/ alkane hydroxylase could ideally be used for dioic acid accumulation, since the dioic acids would not be broken down by β-oxidation. E150 derivatives could be tested for alkylbenzene biotransformations due to convenient assaying using specific TLC plates, since the substrates and degradation products involved are highly UV-active (figure 1.9.).

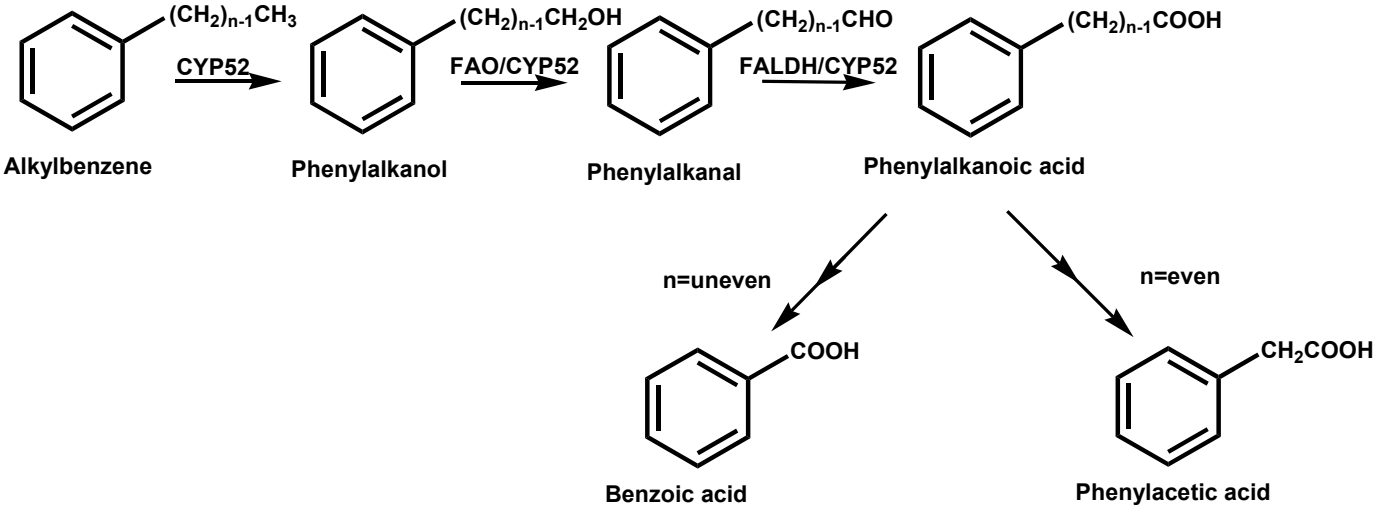


Figure 1.9: Alkylbenzene degradation by *Y. lipolytica* strains with intact β-oxidation pathways.

Eventually, the focus would shift toward investigations of the performance of the self-sufficient P450 CYP102A1 enzymes in the *Y. lipolytica* host system, since the substrate uptake problems encountered with whole-cell biotransformations using this enzyme in *E. coli* as a host would be avoided using *Y. lipolytica* as a host. We obtained the cDNA of CYP102A1 in a plasmid from a group at the University of Edinburgh, and could thence clone this gene under the *POX2* and *ICL* promoters. Both FT-120 and TVN496 strains could be used for expression of this gene, and resultant strains could be tested for fatty acid hydroxylation and alkylbenzene biotransformations, respectively.

The aims of this study were thus to:

(1) Modify the current expression vector systems in our research group to obtain optimal heterologous single and multiple copy expression under the regulation of the p*POX2* and p*ICL* promoters. The promoters would then be compared for expression of the CYP53 and CYP557 enzymes previously cloned in our research group;

(2) Construct strains of *Yarrowia lipolytica*, (with fully intact as well as disrupted β-oxidation pathways), expressing desired P450s under either promoter, as well as more suitable control strains for more thorough comparisons;

(3) Evaluate the transformed strains for product formation and accumulation properties;

(4) Heterologously express the self-sufficient CYP102A1 BM-3 in *Y. lipolytica* using the *POX2* and *ICL* systems for activity evaluation.

Chapter 2: Materials and Methods

Part A: General materials

2.1. Enzymes, kits, general chemicals and reagents

Restriction endonucleases, T4 DNA ligase, Klenow fragment, molecular weight markers and λ DNA (used for preparing *Eco*RI/*Hin*dIII digested λIII marker) were supplied by Fermentas.

The GFX^{TM} PCR DNA and gel band purification kit used was obtained from Amersham Biosciences.

Oligonucleotide primers (Table 2.4.) were designed and analysed using the 'Oligo Analyzer' tool on the Integrated DNA Technologies (IDT) web page (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/) using default settings. Primers were synthesized and supplied by Inqaba Biotechnical Industries.

Calf Intestinal Alkaline Phosphatase and Expand Long Template PCR reagents were supplied by Roche, while the normal PCR reagents, including Taq polymerase, were obtained from New England Biolabs.

Chemicals were, unless otherwise specified, obtained from Fluka, Merck, or Sigma-Aldrich.

2.2. Strains and vectors

Yarrowia lipolytica strains were stored under liquid nitrogen at -70^{0}C in the MIRCEN yeast culture collection of the University of the Free State (UFS), South Africa. The cultures were frozen in LN broth containing glycerol (7% v/v final concentration). The

yeasts were revived by streaking on YM agar plates supplemented with a vitamin solution. Media are described in the following section. The strains are summarized in table 2.2.

Escherichia coli strain XL-10-Gold [*Tet^r D(mcrA) 183 D(mcrB-hsdSMR-mrr)173end A1 supE44 thi-1 recA1 gyrA96 relA1 lac The [F' proAB lacI^{q}ZDM15 Tn10 (Tet^r) Amy Cam^r ;* Stratagene], was used for plasmid preparations. A dam$^-$ and dcm$^-$ strain of *E. coli* was used for some applications, bearing the following genotypic characteristics: *ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10)TetS endA1 rpsL136 dam13::Tn9 xylA-5 mtl-1 thi-1 mcrB1 hsdR2.*

The JMp vectors used in this study were provided by Dr. J-M Nicaud, Laboratoire Microbiologie et Genetique Moleculaire, Institute National Agronomique Paris-Grignon, France. These were later modified to contain the *CYP53* and *CYP557* genes under transcriptional control of the *POX2* promoter, by Dr. A Shiningavamwe during his PhD studies in our laboratory. The vectors used are described in table 2.1.

Table 2.1: Vectors available for modification

Vector	Backbone vector	Promoter	Cloned gene	Terminator	Marker
JMp5	n.a.	None	None	None	*ura3d1*
JMp21-CPR	n.a.	*ICL*	*CPR*	*ICLt*	*leu2*
JMp62	n.a.	*POX2*	None	*LIP2t*	*ura3d1*
JMp64	n.a.	*POX2*	None	*LIP2t*	*ura3d4*
JMp62-CYP53	JMp62	*POX2*	*CYP53*	*LIP2t*	*ura3d1*
JMp64-CYP53	JMp64	*POX2*	*CYP53*	*LIP2t*	*ura3d4*
JMp62-CYP557	JMp62	*POX2*	*CYP557*	*LIP2t*	*ura3d1*
JMp64-CYP557	JMp64	*POX2*	*CYP557*	*LIP2t*	*ura3d4*

Table 2.2: *Y. lipolytica* strains used during this study

Strain	Genotype	Phenotype
Host strains		
FT-120	*MatA ura3-302 leu2-270 xpr2-322 pox1-6::lox pJMp21::CPR*	Ura$^-$, Leu$^-$, ΔAEP, ΔPOX1-6, pPOX2-CPR
E150	*MatB his1 ura3-302 leu2-270 xpr2-322*	Ura$^-$, Leu$^-$, ΔAEP
TVN496	E150 transformed with JMp21-CPR linearized with *Not*I	Ura$^-$, Leu$^-$, ΔAEP, pICL-CPR
Activity control strains		
TVN91	E150 transformed with JMp64-CYP53, then JMp21-CPR, both linearized with *Not*I	Ura$^-$, Leu$^-$, ΔAEP, pPOX2-CYP53, pICL-CPR
TVN210	TVN496 transformed with JMp62-CYP557 linearized with *Not*I	Ura$^-$, Leu$^-$, ΔAEP, pICL-CPR, pPOX2-CYP557
TVN348	TVN496 transformed with JMp64-CYP557 linearized with *Not*I	Ura$^-$, Leu$^-$, ΔAEP, pICL-CPR, pPOX2-CYP557

A 5 kb segment of *Bacillus megaterium* chromosomal DNA, including the CYP102 BM3 gene, cloned in a pUC119 plasmid (pBM23) was generously provided by Professor S.K. Chapman, University of Edinburgh, U.K.

The pGEM-T Easy vector used to clone PCR products was obtained from Promega.

Plasmid maps were constructed and modified using the pDRAW32 program version 1.1.93 of ACACLONE software, based on information obtained from NCBI databases (http://www.ncbi.nlm.nih.gov/).

2.3. Cultivation media and conditions

YPD broth contained, per litre of distilled water, 10 g yeast extract, 20 g peptone and 20 g glucose. For YPD agar plates the media was supplemented with 2g agar per litre distilled water. LN broth contained, per litre of distilled water, 40 g glucose, 10 g tryptone and 10 g yeast nitrogen base (YNB) complete with amino acids and ammonium sulphate. YM agar plates contained 10 g glucose, 10 g peptone, 10 g malt extract, 2 g yeast extract and 2 g agar, per liter of distilled water. YNB selection media contained, per liter distilled water, 1.7 g yeast nitrogen base (YNB) without ammonium sulphate and amino acids, 4 g ammonium chloride, 10 g glucose, 2 g casamino acids (DIFCO Laboratories) and 2 g agar. The pH of the media was adjusted to 6.5 using sodium hydroxide.

Plasmid carrying *E. coli* strains were grown in Luria – Bertani (LB) media, which contained, per litre distilled water, 5 g yeast extract, 10 g sodium chloride, and 10 g tryptone; or on LB plates with additional 2 g agar per litre water. The LB plates and media were supplemented with 50 µg / ml Kanamycin or 100 µg / ml Ampicillin to maintain selective pressure.

Yarrowia lipolytica strains were cultivated at 28^{o}C in Erlenmeyer flasks on a rotary shaker with a speed of 180 rpm. The cultivation broth occupied a volume equal to one tenth of the maximum volumetric capacity of the flask to maximize oxygen transfer. Plates containing *Y. lipolytica* strains were also incubated at 28^{0}C until use, after which they were refrigerated. Yeast strains were cryopreserved at -70^{0}C in YPD media containing 20-50% final concentration glycerol.

E. coli cells were grown in 5 ml LB media in 25 ml test tubes on a rotary shaker at 37^{0}C. *E. coli* containing LB plates were also incubated at 37^{0}C until use, after which they were refrigerated.

Part B: Molecular techniques

2.4. Construction and modification of expression vectors

2.4.1. PCR amplification of genes/inserts

Polymerase chain reactions were performed using either the Taq polymerase or the Expand Long Template System (Roche) according to the specifications of the manufacturers. The reaction mixtures had the following composition: 5 µl (10%) of the appropriate 10 x concentrated buffer system (generally buffer number 3 in the case of the Expand Long Template System), 1 µM of each of the forward and reverse primers, 200 µM dNTP mix, 0.5 µg of template DNA and 3.75 U of appropriate polymerase mix. The reaction mixtures were filled to 50 µl with PCR grade distilled water. Hot-start PCRs were performed in an Applied Biosystems Thermal Cycler 2720 according to standard PCR settings (Table 2.3.). The primers used for various purposes in this study are listed in table 2.4.

Table 2.3: Standard PCR profiles used in this study

Step	Temperature	Step duration		Number of cycles
		Taq	Expand	
Initial denaturation	94°C	2 min	2 min	1
Denaturation	94°C	15 sec	10 sec	
Annealing	55°C[a]	30 sec	30 sec	10
Elongation	72°C	75 sec[b]	4 min[c]	
Denaturation	94°C	15 sec	10 sec	
Annealing	55°C[a]	30 sec	30 sec	
Elongation	72°C	75 sec + 5 sec for each successive cycle	4 min + 20 sec for each successive cycle	15
Final elongation	72°C	7 min	7 min	1
Initial denaturation	94°C	2 min		

[a] 55°C given as an example, annealing temperature was determined by the melting temperatures of the specific primers used (2-5 °C lower than primer Tm).

[b] Elongation time was based on 45 sec/ kb.

[c] Elongation time was dependant on the expected amplicon size.

Table 2.4: List of oligonucleotide primers used in this study and their intended amplicons. Introduced restriction sites are underlined and in bold.

Primer name	Sequence (5' – 3')[a]	Restriction sites introduced	Amplicon
pICLEcoRIR	GAATTCGGTACCGGGCCCGTC	None	*ICL*
pICLMluHpaF	C**ACGCGTTAAC**CTGGGTTAGTACGGGACA	*Mlu*I and *Hpa*I	*ICL*
53+LIP2tMluIF2	C**ACGCGT**CCTAGGATGGGCATAG	*Mlu*I	*CYP53* and *LIP2t*
53+LIP2tClaIR2	C**ATCGAT**GAATTCGATTTGTCTTAGAGGAACG	*Cla*I	*CYP53* and *LIP2t*
p53 F1	G**ACGCGT**AATCACACAAGCAACGGATCC	*Mlu*I	*CYP53*
p53R1	GCATCGATACCACAGACACCCTAGG	None	*CYP53*
C557HpaIF	**GTTAAC**ATGCTCGCGCTCGTCTGC	*Hpa*I	*CYP557*
557AvrIIR3	**CCTAGG**CTCATCGTCGGGATATTG	*Avr*II	*CYP557*
BM3MluIF	C**ACGCGT**ATGACAATTAAAGAAATGCCTCAGC	*Mlu*I	*CYP102A1*
Bm3R1	**CCTAGG**TTACCCAGCCCACACGTCTTC	*Avr*II	*CYP102A1*
Bm3F2BamHI	CCCGGC**GGATCC**ATGACAATTAAAGAAATGCC	*Bam*HI	*CYP102A1*

[a]Introduced restriction sites are indicated in bold face

2.4.2. Plasmid digestions and ligations

Restriction digests of plasmid DNA were performed according to the specifications of the manufacturers. The appropriate ratios of restriction endonucleases and most suitable buffer systems required for optimal double digestions were obtained on the Fermentas website in the 'Double digest' section (www.fermentas.com/doubledigest/index.html).

Ligation reactions were carried usually with a 4:1 insert: backbone ratio. When blunt end ligations were carried out, the vector backbones were subjected to dephosphorylation using alkaline phosphatase (AP) from calf intestine to prevent re-ligation of the backbone on itself. The DNA was incubated with 2 U of AP for 15 minutes at 37^{o}C, followed by further incubation at 55^{o}C for 45 minutes, before another

2 U AP was added and incubated for 60 minutes at 37^{o}C. The dephosphorylated DNA was then purified.

Klenow treatment of DNA was performed using 25 U of Klenow fragment in supplied buffer and 50 µM dNTP mixture for 30 minutes at 37^{o}C, after which it was purified. The resulting blunt ends of the DNA were then re-ligated using ligase following the manufacturer's instructions.

2.4.3. Visualization and purification of PCR and digestion products

PCR amplicons and restriction digest mixtures were loaded into 1% (w/v) agarose gels containing 2.5 mg/ml ethidium bromide. The TAE buffer (pH 8.5) used for the preparation of the gels and as the electrophoresis buffer contained 40 mM Tris, 2 mM EDTA and 20 mM glacial acetic acid. The gels were electrophoresed at 90 V in TAE buffer (40 mM Tris, 2 mM EDTA, 20 mM glacial acetic acid, pH 8.5) until sufficient band migration had occurred. The gels were visualised using a ChemiDoc XRS system (BioRad Laboratories) under short wavelength UV light allowed capturing and storing of the gel profile.
Gels were visualised using a DarkReader^{TM} transilluminator (Fermentas) allowing excision of bands for further cloning. DNA excised from the gel were purified and eluted using 50 µl of a 10 mM Tris-HCl solution (pH 8.0). DNA concentrations were determined using an Eppendorf Biophotometer.

2.4.4. Transformation of *E. coli* and mini-preparation of vector DNA

All PCR amplicons were sub-cloned into pGEM$^{®}$-T Easy according to the manufacturer's instructions, before further application. The pGEM$^{®}$-T Easy vector containing the inserted amplicon was used to transform 50 µl aliquots of *E. coli* cells which had been made competent using a modified version of the $RbCl_2$ method originally described by Hanahan (1983). The transformed cells were streaked onto LB plates supplemented with the appropriate antibiotic (section 2.3.) as well as 10 mg/ml

IPTG (isopropyl-*β*-D-1-thiogalactopyranoside) and 40 mg/ml X-gal (5-bromo-4-chloro-3-indolyl- *β*-D-galactopyranoside) for selection purposes. Plates were incubated at 37^{o}C for 16 hours. Positive transformants were selected and used to inoculate 5 ml LB media containing appropriate antibiotic, and incubated for 16 hours at 37 ^{o}C.

Inserts removed from vectors via restriction digestion were ligated into the backbones of the desired plasmids, and the resultant plasmid was used to transform *E. coli* by the same method described above. Transformants were screened for the desired insert by means of colony PCR. Transformants showing the correct size PCR product were used to inoculate 5 ml LB media containing the appropriate antibiotic, and incubated for 16 hours at 37 ^{o}C.

DNA mini-preparations were performed on the 5 ml LB grown transformant cells using the lysis by boiling method of Sambrook *et al*. (1989). As the final step the precipitated DNA was resuspended in TE buffer containing 50 µg/ml RNase and incubated at 37^{o}C for 1 hour. Thereafter mini-preparations were screened for positive recombinant plasmids using restriction analysis.

2.4.5. Modification of vectors and amplification of cytochrome P450 genes

Purified vectors and insert DNA samples were stored at -20^{o}C until required. The JMp5 vector was used as a starting vector, and the vectors JMp21-CPR and JMp64 were the sources of the *ICL* promoter and the defective *ura3d4* marker, respectively (table 2.1.). The cDNA of the *CYP53* and *CYP557* genes were available in the JMp62 (single-copy integration) and JMp64 (multi-copy integration) vectors, while the CYP102A1 cDNA was available in the pBM23 vector.

The *CYP53* gene was amplified along with the *LIP2* terminator region using the primers 53+LIP2tMluIF2 and 53+LIP2tClaIR2. The primers C557HpaIF and 557AvrIIR3 were designed for the amplification of the CYP557 gene. Primer pair Bm3F2BamHI and BM3R1 were designed to amplify the CYP102A1 gene for fusion to

POX2 promoter and primers BM3MluIF and Bm3R1 were used for amplification of *CYP102A1* for fusion to *ICL* promoter.

2.5. Preparation and transformation of competent *Yarrowia lipolytica* cells

Preparation of competent *Yarrowia lipolytica* cells and transformation with vector DNA was done according to a modification of the method of Barth and Gaillardin (1996). Two day old cells of *Yarrowia lipolytica* strains (tables 2.2.) from an YP_2D_2 plate were used to inoculate 20 ml YP_2D_2 media supplemented with 50 mM citrate buffer as a pre-culture and incubated at 28^{o}C for 15 hours on a rotary shaker (180 rpm). When an OD_{600} of 5 was reached, cells were used to inoculate another 20 ml YP_2D_2 culture supplemented with 50 mM citrate buffer (pH 4) and incubated at 28^{o}C on a rotary shaker (180 rpm). When a concentration of 10^8 cells/ml was reached, cells were harvested via centrifugation for 10 minutes at 4000 *g* in a Megafuge 1.0R centrifuge (Heraeus Instruments). Cells were rinsed twice with TE buffer (10 mM Tris, 1 mM EDTA, pH 8), before being resuspended in 0.1 M LiAc (pH 6) to a final concentration of $5x10^7$ cells/ml. The cells were transferred to a sterile 50 ml flask and incubated at 28^{o}C for 1 hour with gentle shaking. The cells were then centrifuged for 5 minutes at 2500 *g*. The competent cells were resuspended in LiAc to a final concentration of $5x10^8$ cells/ml, before being aliquoted in 100 µl aliquots in 2 ml microcentrifuge tubes.

5 µl carrier DNA (5 mg/ml salmon sperm DNA in 50 mM Tris, 5 mM EDTA) and 10 µl of *Not*I digested vector DNA were added to the 100 µl competent cells with gentle mixing. The transformation mix was incubated in a 28^{o}C water bath for 15 minutes. After the incubation, 700 µl of 40% PEG_{4000} in 0.1 M LiAc (pH 6) was added before incubation for an hour at 28^{o}C with gentle shaking. Cells were heat shocked for 10 minutes at 39^{o}C before 1.2 ml of 0.1 M LiAc (pH 6) was added. The cells were then centrifuged at 2500 *g* for 5 minutes in an Eppendorf Centrifuge 5415D. Cells were resuspended in 200 µl LiAc and plated on YNB plates using glass beads.

Plates for single copy transformants were incubated for 3-5 days before colonies appeared, while plates for multi-copy transformants required 10 to 15 days incubation before transformants appeared.

2.6. Confirmation of genomic integration of heterologous cytochrome P450 genes

Transformants were subjected to PCR screening for the presence of the cloned P450 of interest. The reactions were performed either on whole cells or on isolated genomic DNA. The PCR profile used for the confirmation is described in section 3.4.1.

Total genomic DNA was isolated from 2 ml of 24 hour old cultures of transformed *Y. lipolytica* strains. The strains were harvested for 1 minute at 13000 rpm and resuspended in 500 µl lysis buffer [100 mM Tris (pH 8); 50 mM EDTA; 1% Sodium Dodecyl Sulphate (SDS)] and 200 µl acid washed glass beads (400-500 microns) by vigorous vortexing for 4 minutes. Broken cells were immediately cooled on ice for 5 minutes. The supernatant was salvaged and 275 µl of 7 M ammonium acetate (pH 7) was added. The mixture was vortexed, incubated for 5 minutes at 65^{o}C and cooled on ice for 5 minutes. Chloroform (500 µl) was added, and the mixture was vortexed, centrifuged at 14000 rpm at 4 ^{o}C for 2 minutes in an Eppendorf Centrifuge 5417R. The supernatant was subjected to DNA precipitation using an equal amount of isopropanol for 30 minutes at -20 ^{o}C. The tubes were centrifuged again at 14000 rpm at 4^{o}C for 2 minutes, and the pellet was washed with 70% cold ethanol. The pellet was dried and dissolved in 50-100 µl TE buffer (pH 8) supplemented with 50 µg / ml RNase and incubated at 37^{o}C for 1 hour. Isolation of genomic DNA was confirmed on an agarose gel.

For colony PCR, cells from transformation plates were resuspended in distilled water (20 µl) and incubated at 94 ^{o}C for 10 minutes. Following this incubation step, the remainder of the amplification mixture was added and the standard PCR procedure was followed.

2.7. Sequence analysis of DNA fragments

DNA sequencing was performed by Inqaba Biotec South Africa on DNA prepared according to their specifications. Sequence editing and analysis was performed using the Chromas version 2.13 software of Technelysium (Pty Ltd). Sequence alignments were performed using the ClustalW alignment tool of the European Bioinformatics Institute (EBI) at URL: http://www.ebi.ac.uk/clustalw/index.html.

Part C: Biotransformations using recombinant strains

2.8. General whole cell biotransformation procedure

Pre-cultures in YPD broth (10 ml in 100 ml flasks) were inoculated from cultures grown for two days on YPD plates. Main cultures in YPD broth containing 50 mM potassium phosphate buffer (10 ml in 100 ml flasks) were inoculated from with 10% v/v of 24 hour old pre-cultures. All cultures were incubated at 28^0C on a rotary shaker (180 rpm).

Main cultures were incubated for 24 hours before inducers were added. After allocated induction times, substrates were added. Inducers were sometimes supplemented at regular intervals after initial induction (see section 2.10).

When growth was monitored, cells were harvested from representative samples (500 µl) by centrifugation at 9300 x *g* for 10 minutes. The harvested cells were appropriately diluted in physiological salt solution (0.9% w/v NaCl) to a final volume of 200 µl for determination of the optical density. Absorbance readings were read on a Labsystems iEMS reader MF using the Ascent software for iEMS applications (Thermo BioAnalysis Company, Helsinki, Finland).

2.9. Sample extraction and analysis

After substrate addition, samples (500 µl) were taken at regular intervals and acidified using hydrochloric acid (5M, 70 µl) to a pH below 3. Ethyl acetate (300 µl) containing myristic acid (4.38 mM), as an internal standard, was added to the samples. After thorough mixing by inversion, samples were centrifuged for 10 minutes at 9 300 x *g*. The upper organic layer was collected and used for further analysis.

Analysis was done using thin layer chromatography (TLC) and gas chromatography (GC). For TLC aliquots (5-20 µl) of organic extracts of samples and standards were spotted on Alugram® silica gel F_{254} TLC plates (Merck) and the plates were developed using a mobile phase containing di-n-butyl ether/formic acid/distilled water (90:7:3 v/v). The plates were dried and viewed under UV light. Aromatic compounds appeared as UV absorbing spots.

For visualization of non-UV active acids, plate staining was performed. The plates were heated for 15-20 minutes (110^{o}C), immersed in Bromocresol green stain (0.1g Bromocresol green, 500 ml ethanol and 5ml 0.1M NaOH), dried and heated again for 15-20 minutes (110^{o}C). Acids appeared as yellow spots on a blue-green background.

Organic acids were methylated prior to GC analysis using equal volumes of a trimethylsulfonium hydroxide (TMSH) preparation (Butte, 1983). GC analyses were done on samples (1µl) using a Hewlett-Packard 5890 series II gas chromatograph equipped with a 30 M x 0.53mm Chrompack® CP wax 52 CB column, with carrier gas, H_2, at 5ml/min., split ratio of 1:40, inlet temperature at 200^{o}C, initial temperature at 120^{o}C for 5 min, increased at 10^{o}C/min to 250^{o}C for 30 minutes. Flame ionisation detector (FID) temperature was at 300^{o}C.

2.10. Comparative biotransformations using recombinant strains

The initial screening for the respective activities of the transformants was performed, from which a selection of the best performing strains was made for further comparative studies. The experimental outlay for the biotranformations is given in table 2.5.

In the case of experiment 7, in which alkane hydroxylases were screened for dodecane dioic acid accumulation from dodecane, the pH of each culture was monitored using phenol red (2.5 mg / ml phenol red; 10mM NaOH) in a final concentration of 8 µl / ml, and adjusted to 8 when necessary.

Stock solutions of the fatty acids and the *p*-nonyloxybenzoic acid contained Tween 80 to assist emulsification. The final concentration of Tween 80 in the biotransformation media was 0.1 % in these cases.

Table 2.25: Experimental outlay for the biotransformation experiments.

Exp.	Hydroxylase activity	Test strains	Induction	Substrate
1	Benzoate *para*-hydroxylase	TVN496 transformants with *ICL*tr:*CYP53* TVN496 transformants with *POX2:CYP53*	Ethanol (2 % v/v; 434 mM) after 24 h growth Oleic acid (2 % v/v; 71 mM) after 24 h growth	Benzoate (0.5 v/v; 41 mM) after 48 h growth
2	Benzoate *para*-hydroxylase	TVN496 transformants with *ICL:CYP53* FT120 transformants with *ICL:CYP53*	Ethanol (2 % v/v; 434 mM) after 24 h growth	Benzoate (0.5 v/v; 41 mM) after 48 h growth
		CTY004; CTY006 FT120 transformants with *POX2:CYP53*	Oleic acid (2 % v/v; 71 mM) after 24 h growth	
3	Benzoate *para*-hydroxylase	CTY004; CTY006; CTY007; TVN91; CTY012	Ethanol (2 % v/v; 434 mM) after 24 h, ethanol (1 % v/v; 217 mM) & oleic acid (1 % w/v; 35 mM) after 48 h; ethanol (0.5 % v/v; 109 mM) after 72 h growth	Benzoate (0.5 v/v; 41 mM) after 30 h growth
			Oleic acid (2 % v/v; 71 mM) after 24 h oleic acid (1 % w/v; 35 mM) after 48 h; oleic acid (0.5 % v/v; 18 mM) after 72 h growth	
4	Alkane hydroxylase	TVN496 transformants with *ICL:CYP557*; CTY004;	Ethanol (2 % v/v; 434 mM) after 24 h growth	Hexylbenzene (1 % v/v; 62 mM) after 48 h growth
		TVN210	Oleic acid (2 % v/v; 71 mM) after 24 h growth	
		TVN348	Ethanol (2 % v/v; 434 mM) & oleic acid (2 % v/v; 71 mM) after 24 h growth	
5	Alkane hydroxylase	CTY004, CTY013; CTY014	Ethanol (2 % v/v; 434 mM) after 24 h, ethanol (1 % v/v; 217 mM) after 48 h; ethanol (0.5 % v/v; 109 mM) after 72 h growth Oleic acid (2 % v/v; 71 mM) after 24 h; oleic acid (1 % w/v; 35 mM) after 48 h; oleic acid (0.5 % v/v; 18 mM) after 72 h growth	Hexylbenzene (1 % v/v; 62 mM) after 30 h growth
6	Alkane and fatty acid hydroxylase	FT120 transformants with *ICL:CYP557*	Ethanol (2 % v/v; 434 mM) after 24 h growth	Dodecane (1 % v/v; 59 mM) after 48 h growth
7	Alkane and fatty acid hydroxylase	CTY027; CTY028; CTY029	Ethanol (0.5 % v/v; 109 mM) after 24, 48 and 72 h growth	Dodecane (1 % v/v; 59 mM) after 24 h growth
			Glucose (0.5 % v/v; 28 mM) after 24, 48 and 72 h growth	
8	Subterminal fatty acid hydroxylase	CTY004; CTY018; CTY019	Ethanol (2 % v/v; 434 mM) after 24 h, ethanol (1 % v/v; 217 mM) after 48 h; ethanol (0.5 % v/v; 109 mM) after 72 h growth	4-Nonyloxybenzoic acid (0.1 % w/v; 4 mM) after 30 h growth
			Oleic acid (2 % v/v; 71 mM) after 24 h; oleic acid (1 % w/v; 35 mM) after 48 h; oleic acid (0.5 % v/v; 18 mM) after 72 h growth	
9	Subterminal fatty acid hydroxylase	CTY029; CTY030	Ethanol (2 % v/v; 434 mM) after 24 h, ethanol (1 % v/v; 217 mM) after 48 h; Ethanol (0.5 % v/v; 109 mM) after 72 h growth	Myristic acid (1 % w/v; 22 mM) after 30h
				Palmitic acid (1 % w/v; 20 mM) after 30h

Chapter 3: Results and Discussion

This study was aimed at improving the expression system currently used in our research group for the heterologous expression of cytochrome P450 genes. Previous work conducted in our group found limitations in P450 expression when the *POX2* promoter was used, notably the unintentional induction of native P450s within the cells (Shiningavamwe, 2004; Van Rooyen, 2005; Obiero, 2006). The *ICL* promoter was suggested as a possible alternative, since it can be induced using either substances not compatible with native P450s (i.e. ethanol), or substances that do induce native P450s (i.e. fatty acids) (Shiningavamwe *et* al, 2006).

The primary goal of this research project was to construct vectors which would allow expression of cloned P450s (mono- or multi-copy) under the control of the *ICL* promoter, for comparison with genes under control of the *POX2* promoter. Ideally, the presence of ethanol would simultaneously induce the *ICL* promoter and inhibit the native P450s in the *Y. lipolytica* cells, allowing determination of the cloned P450 activity only.

The constructed vectors would be used to transform strains of *Y. lipolytica*. The eexpression systems would be tested using the gene coding for benzoate-*para*-hydroxylase, *CYP53B1,* as a reporter gene, followed by determination of *CYP557* activity, and ultimately for cloning of the self-sufficient *CYP102A1* into *Y. lipolytica*. Additionally, suitable control strains were to be constructed through the cloning of parent vectors devoid of any expression cassette inserts.

Molecular techniques

3.1. Construction and modification of expression vectors

The first step was to ensure that there were versions of the plasmid allowing either mono-copy or multi-copy integration into the host genome. The defective allele of the Ura marker, ura3d4, was removed from JMp64 using *Cla*I and *Stu*I, and used to replace the non-defective allele ura3d1 in JMp5, which was excised using the same restriction enzymes. The resultant ligation produced vector CMp1a (figure 3.1).

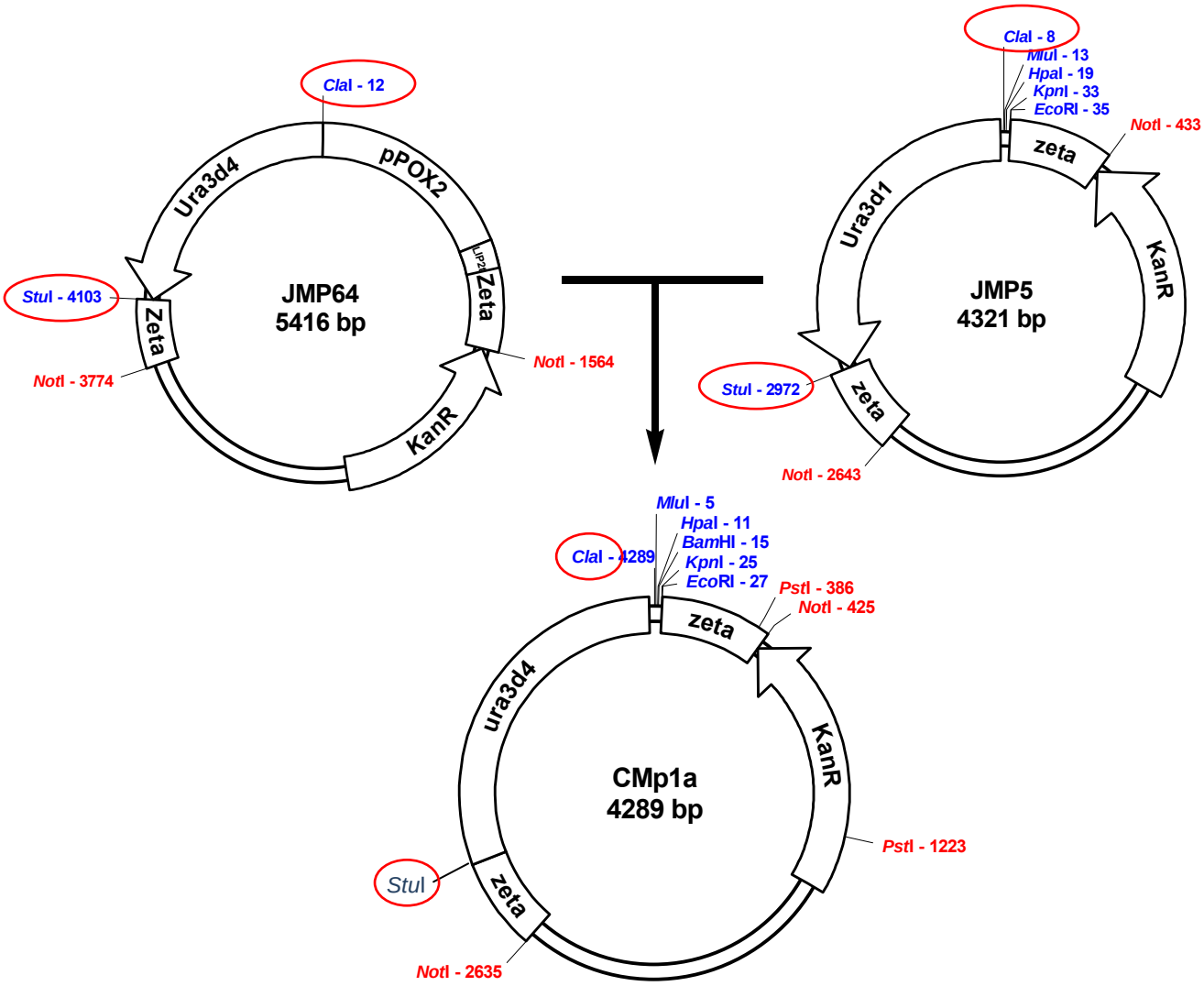


Figure 3.1: Replacement of the monocopy marker (ura3d1) in JMp5 with the multi-copy marker (ura3d4) by restriction with *Cla*I and *Stu*I, producing CMp1a. The *Stu*I and *Cla*I sites used for these steps are encircled in all vectors.

3.1.1. Variations of the expression cassette

A set of vectors, CMp2a-e with truncated versions of the *ICL* promoter were constructed. The 2536 bp region upstream of the *Y. lipolytica* isocitrate lyase (ICL) gene contains the gene's promoter, and an intron between the adenosine and thymidine bases of the ATG start codon of the *ICL* gene (figure 3.2). Nthangeni *et al* (2004) had cloned this entire region, up to the end of the intron, into the JMp21 vector, where it was fused to the P450 reductase (CPR) gene of *Y. lipolytica.*

gTATATATAg atctggG'GAT CCccagtaga ctgaccaagc atacaaaaAg tgagtatcca acagcgacac gtgagatggc agagacacag agacgtgtct acatggttgg acaagtctcc acattcgcca gagacgtatc cacatacaaa cacaatctca cagctgatct gctcctgtga cagcacagta catgttagtg gatgaggtgt tgtgtagtgg gttaaatggg tggactgatt cagtggcatc ggtggcgaca ccctctactc ttcatgtcgt cacctaccgt tcggaatccc aattatctga tgaactaaac gatttctggc caaaacacaa ttttgccaaa gaagtcggtc tcaccaatgc aagtgtcaca tcaaacatct gtcccgtact aacccagTGt

Figure 3.2: Upstream region of the *Y. lipolytica* isocitrate lyase (ICL) gene, from base pair 2131 to 2540. The first highlighted region represents the TATA box of the promoter. The second is a *Bam*HI site available for cloning. The remaining two highlighted sections represent the ATG start codon for the isocitrate lyase gene, interrupted by an intronic region.

The intron apparently serves no particular purpose. Indeed, Mauersberger and colleagues achieved the same expression with or without this intronic region (Mauersberger, personal communication). Also, we had no interest in the start codon of the *ICL*, since the intention of this study was fusing the promoter to the cDNA of different genes with their own start codons.

The presence of the *Bam*HI recognition site (figure 3.2, second highlighted region) was an attractive option for our cloning strategy. The 32 bp region preceding the adenosine nucleotide of the *ICL* start codon, especially the 5

bases immediately preceding the adenosine, was however presumed potentially crucial for effective gene expression (Mauersberger, personal communication). If the *Bam*HI site was used for cloning of inserts, those 32 bp would be lost. The TATA box (first highlighted region) would however survive restriction using this enzyme.

Despite the risk, it was initially decided to venture on convenience and use the *Bam*HI restriction site, losing the region following it. A truncated part of the *ICL* promoter was removed from the JMp21-CPR vector using *Eco*RI and *Bam*HI, and cloned into the vector JMp5, resulting in vector CMp2a (figure 3.3).

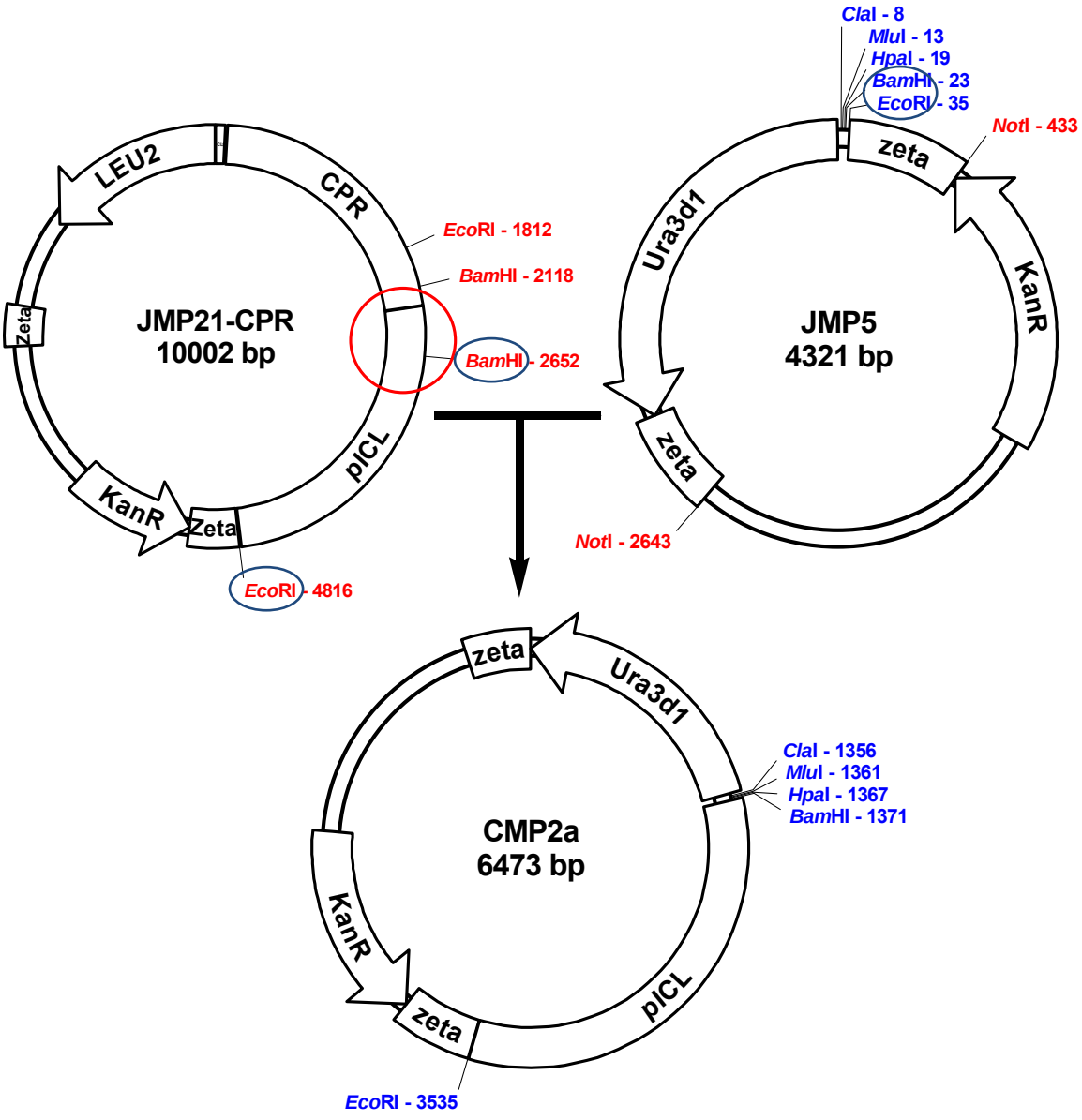


Figure 3.3: Construction of the CMp2a vector, by ligation of a truncated ICL promoter (ICLtr) from JMp21, released by restriction with *Eco*RI and *Bam*HI, into JMp5. This plasmid served as the vector backbone for CMp2b-d. The *Eco*RI and *Bam*HI sites used for the cloning of *ICLtr* are encircled in all vectors, and the region omitted by using *Bam*HI is also encircled in the parent vector.

The *CYP53* gene and *LIP2*t terminator from JMp62-CYP53 were used to test the effectiveness of the truncated *ICL* promoter. A PCR strategy was devised to test the necessity of the region lost through the use of the *Bam*HI site,

replacing some of the nucleotides numerically with the last 25 bp of the *POX2* promoter. This could test whether the effect of these bases, if they were essential, could be attributed to the distance they impart between the promoter and the start codon, rather than the sequence itself. An alignment of the 3' sequences of the *ICLtr* promoters used in this investigation, as compared to the 3' end of the complete promoter, is provided in figure 3.4. The alignment shows where the section of *POX2* was incorporated, and also illustrates the difference between the complete *ICL* promoter and *ICLtr* at the 3' end. The site where the *CYP53* gene was fused to each version of the promoter is also indicated.

```
ICLtr/POX       TATATATAGATCTGGGGATCCGTTAACGCGTAATCACACAAGCAACGGATCCGGTACCTA 60
ICLtr           TATATATAGATCTGGGGATCCGTTAACGCGTCCTAGGATGGGC---------------- 43
ICL             TATATATAGATCTGGGGATCCCCAGTAGACTGACCAAGCATACAAAAAGTG-AGTATCCA 59
                ********************      *  *           *
ICLtr/POX       GGATGGGC------------------------------------------------------ 68
ICLtr           -------------------------------------------------------------
ICL              ACAGCGACACGTGAGATGGCAGAGACACAGAGACGTGTCTACATGGTTGGACAAGTCTCC 119

ICLtr/POX       -------------------------------------------------------------
ICLtr           -------------------------------------------------------------
ICL              ACATTCGCCAGAGACGTATCCACATACAAACACAATCTCACAGCTGATCTGCTCCTGTGA 179

ICLtr/POX       -------------------------------------------------------------
ICLtr           -------------------------------------------------------------
ICL             CAGCACAGTACATGTTAGTGGATGAGGTGTTGTGTAGTGGGTTAAATGGGTGGACTGATT 239

ICLtr/POX       -------------------------------------------------------------
ICLtr           -------------------------------------------------------------
ICL             CAGTGGCATCGGTGGCGACACCCTCTACTCTTCATGTCGTCACCTACCGTTCGGAATCCC 299

ICLtr/POX       -------------------------------------------------------------
ICLtr           -------------------------------------------------------------
ICL             AATTATCTGATGAACTAAACGATTTCTGGCCAAAACACAATTTTGCCAAAGAAGTCGGTC 359

ICLtr/POX       -------------------------------------------------------------
ICLtr           -------------------------------------------------------------
ICL             TCACCAATGCAAGTGTCACATCAAACATCTGTCCCGTACTAACCCAGACGCGTCCTAGGA 419

ICLtr/POX       -----
ICLtr           -----
ICL             TGGGC 424
```

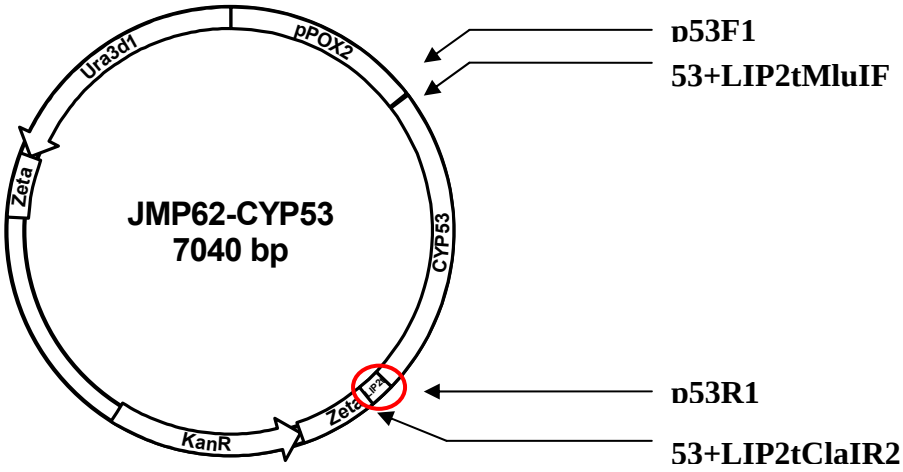
Figure 3.4: ClustalW2 alignment of the different forms of the *ICL* promoter used in this study, fused to the *CYP53* gene.

ICL - complete *ICL* promoter; ICLtr - ICL truncated by 403 bp; ICLtr/POX2 - hybrid of *ICLtr* and *POX2* promoters, containing 25 bp of the *POX2* promoter. Nucleotides colored purple represent sections of the *ICL* promoter. Red nucleotides represent the section of the *POX2* promoter incorporated. Green nucleotides represent the beginning of the *CYP53* cDNA.

*Mlu*I was used for cloning of inserts into vectors containing any form of the *ICL* promoter, while *Avr*II was used for cloning of *CYP53* into the initial *POX2* containing vector. Their recognition sequences (ACGCGT for *Mlu*I and CCTAGG for *Avr*II) indicated by light blue nucleotides.

In order to test the necessity of the terminator for the expression cassette, vectors without external terminator regions were also constructed for testing, thereby relying solely on the stop codon of the expressed gene.
Primers p53 F1, 53+LIP2tMluIF2, p53R1 and 53+LIP2tClaIR2 (table 2.4; figure 3.5.) were designed for the amplification of the reporter gene *CYP53* (with and without the *LIP2*t terminator) from the JMp62-CYP53 vector, while introducing 5' *Mlu*I and 3' *Cla*I sites. The two sets of primers could be combined in four ways to allow four variations of the expression cassette for testing, as illustrated in figure 3.5.

a



b

Vector	Primer combination	Phenotype	*POX2* region	Terminator
CMp2b	A: p53F1 and p53R1	ICLtr/POX2-CYP53	Present	Absent
CMp2c	B: 53+LIP2tMluIF2 and 53+LIP2tClaIR2	ICLtr-CYP53-term	Absent	Present
CMp2d	C: p53F1 and 53+ClaIR2	ICLtr/POX2-CYP53-term	Present	Present
CMp2e	D: 53+LIP2tMluIF2 and p53R1	ICLtr- CYP53	Absent	Absent

c

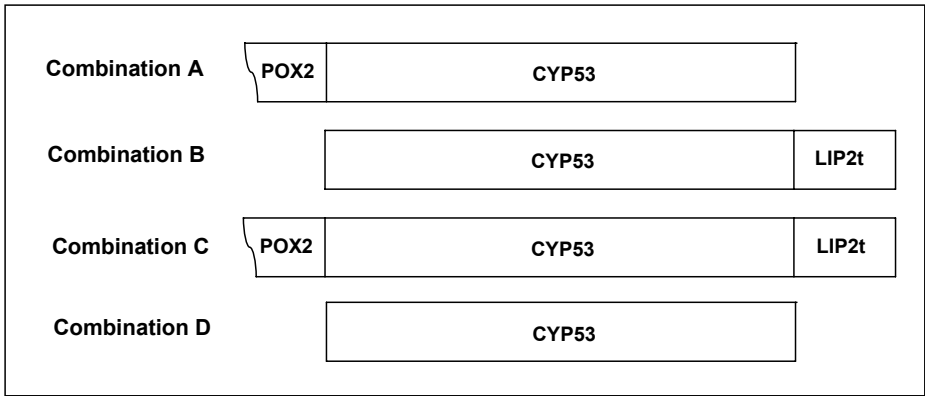


Figure 3.5: (a) Primer binding sites for expression cassette variations, illustrating that the *LIP2*t terminator region (encircled), as well as the last

25 bp of the *POX2* promoter, would be either included or excluded, depending on the primers used. (b) Description of the outcomes of different primer combinations. (c) Illustrative comparison of the PCR products from each primer combination, for fusion to the ICLtr promoter. In ICLtr/POX2-CYP53 (from combination 'A') *CYP53* expression was under the control of the hybrid *ICLtr/POX2* promoter, in the absence of the *LIP2t* terminator. In ICLtr-CYP53-term (from combination 'B') *CYP53* expression was under the control of the *ICLtr* promoter, in the presence of the *LIP2t* terminator. In ICLtr/POX2-CYP53-term (from combination 'C') *CYP53* expression was under the control of the hybrid *ICL/POX2* promoter, in the presence of the *LIP2t* terminator. In ICLtr- CYP53 (from combination 'D') *CYP53* expression was under the control of the I*CLtr* promoter in the absence of the *LIP2t* terminator.

The varying PCR products obtained from the different primer combinations were cloned into CMp2a using *Mlu*I and *Cla*I. Combination A resulted in CMp2b, which contained additional bases at the 5' end (from *POX2* promoter), but no terminator was added. Combination B resulted in CMp2c, which had the terminator added, but no extra bases at the 5' end. Combination C resulted in CMp2d which had both additional 5' bases (from the *POX2* promoter), as well as the terminator. Combination D resulted in CMp2e, which had neither the terminator added nor any extra 5' bases.

Although the two varying factors were only expected to have an impact on the expression level under induction conditions, a surprising observation was made on inspection of the transformation plates. The different combinations also had an effect on the transformation efficiencies of the yeast strains, as can be seen in figure 3.6. These results were repeated for three different strains of *Y. lipolytica* (strains E150, TVN496 and FT-120), and the same pattern was evident in each case. Results are only shown for the FT-120 strain, since it generally has lower transformation efficiency and hence the differences can be observed clearly.

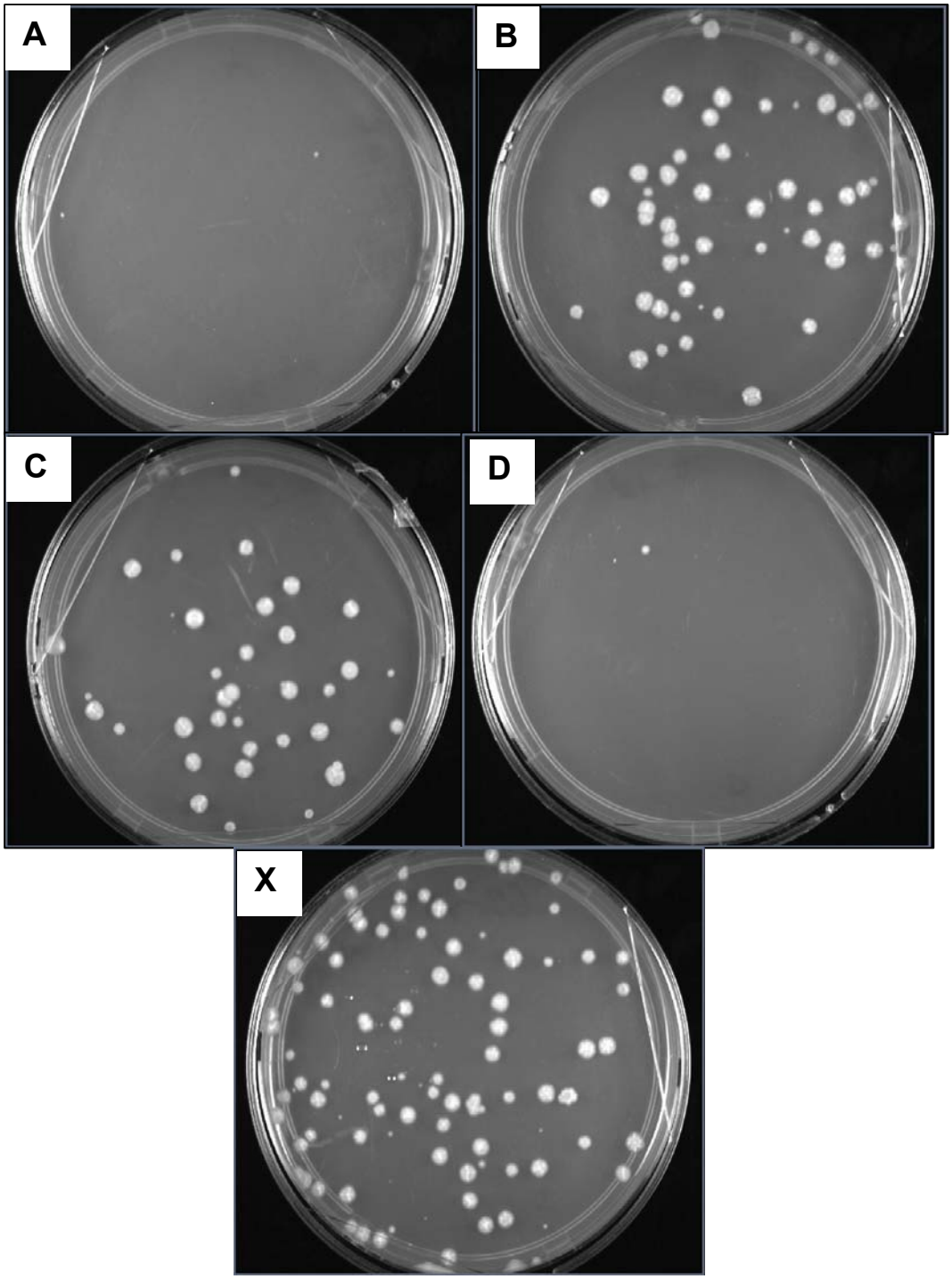


Figure 3.6: Differences in transformation efficiencies of *Y. lipolytica* FT-120 transformed with different variations of the expression cassette. Images A-D represents the results of combinations A-D (figure 3.5) as compared to image X which represents the positive control, which was transformed with the JMp62-CYP53 vector. A – ICLtr/POX2-CYP53; B – ICLtr-CYP53-term; C – ICLtr/POX2-CYP53-term; D – ICLtr-CYP53.

It is clear that the absence of the terminator had a very detrimental effect on the transformation efficiency. The additional 5' bases apparently had very little effect on transformation efficiency.

A few representatives from TVN496 transformants containing CMp2b-e vectors, as well as the control (transformed with theJMp62-CYP53 vector) were selected for activity screening (figure 3.7).

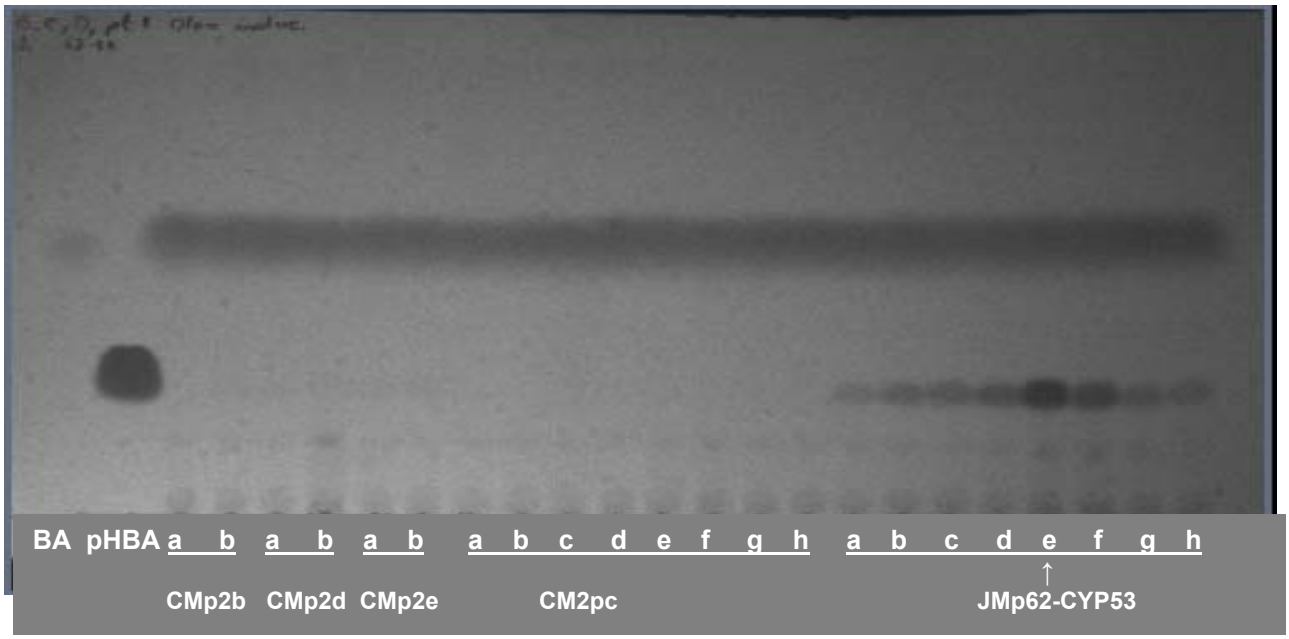


Figure 3.7: TLC plate of benzoate-*para*-hydroxylase activity in screened transformants at 120 hours after substrate addition. BA – benzoic acid, pHBA – *para*-hydroxy benzoic acid, Biotransformations were done with TVN496 transformants containing vector CMp2b (combination A) (two strains), vector CMp2d (combination C) (two strains), vector CMp2e (combination D) (two strains), CMp2c (combination B) (eight strains) and control vector JMp62-CYP53 (eight strains).

From figure 3.7, it is evident that none of these combinations resulted in any visible activity compared to the positive controls even after 120 hours of biotransformation, using an assay technique which is very sensitive to the expected activity. Transformant e (Fig. 3.7 marked with arrow) of the positive

controls, showed promising activity. It was labelled CTY006 and was stored for further use.

Although no benzoate *para*-hydroxylase activity was observed with the strains containing *CYP53* under the modified *ICL* promoter, the low transformation efficiencies observed when the terminator was omitted were still unexpected. It can tentatively be explained as the possibility that the modified *ICL* promoter gave rise to the expression of a faulty protein. Such a situation could lead to the accumulation of unintended products, potentially detrimental to the cells, and hence limiting the number of successful transformants. This especially seems to be the case for the absence of the terminator region, which probably leads to the transcription occurring past the intended point.

Based on these results, we decided to amend the entire strategy and to clone inserts rather under the control of the complete *ICL* promoter.

3.1.2. New strategy for construction of expression vectors

The primers *ICL*MluHpaF and *ICL*EcoRIR (table2.4) were designed to amplify the entire 2535 bp *ICL* promoter from JMp21-CPR, so as to include the restriction sites for *Eco*RI and *Kpn*I at the 5' end which originate from the JMp21 vector, and to introduce 3' *Hpa*I and *Mlu*I restriction sites (part a of figure 3.8). After sub-cloning in pGem-T Easy, the amplicon was removed using *Eco*RI and *Mlu*I, and subsequently ligated into vectors JMp5 for mono-copy integrations and CMp1a for multi-copy integrations. The resultant vectors were named CMp1b (for monocopy integration) and CMp1c (for multi-copy integration) (part a of figure 3.8).

CMp10 was produced by removing the expression cassette of JMp21-CPR using *Mlu*I and *Kpn*I (part b of figure 3.8). The backbone remaining after this excision was treated with Klenow fragment as discussed before (section 2.4.2.) before re-ligation of the resulting blunt ends. This vector would later be used for the generation of control strains.

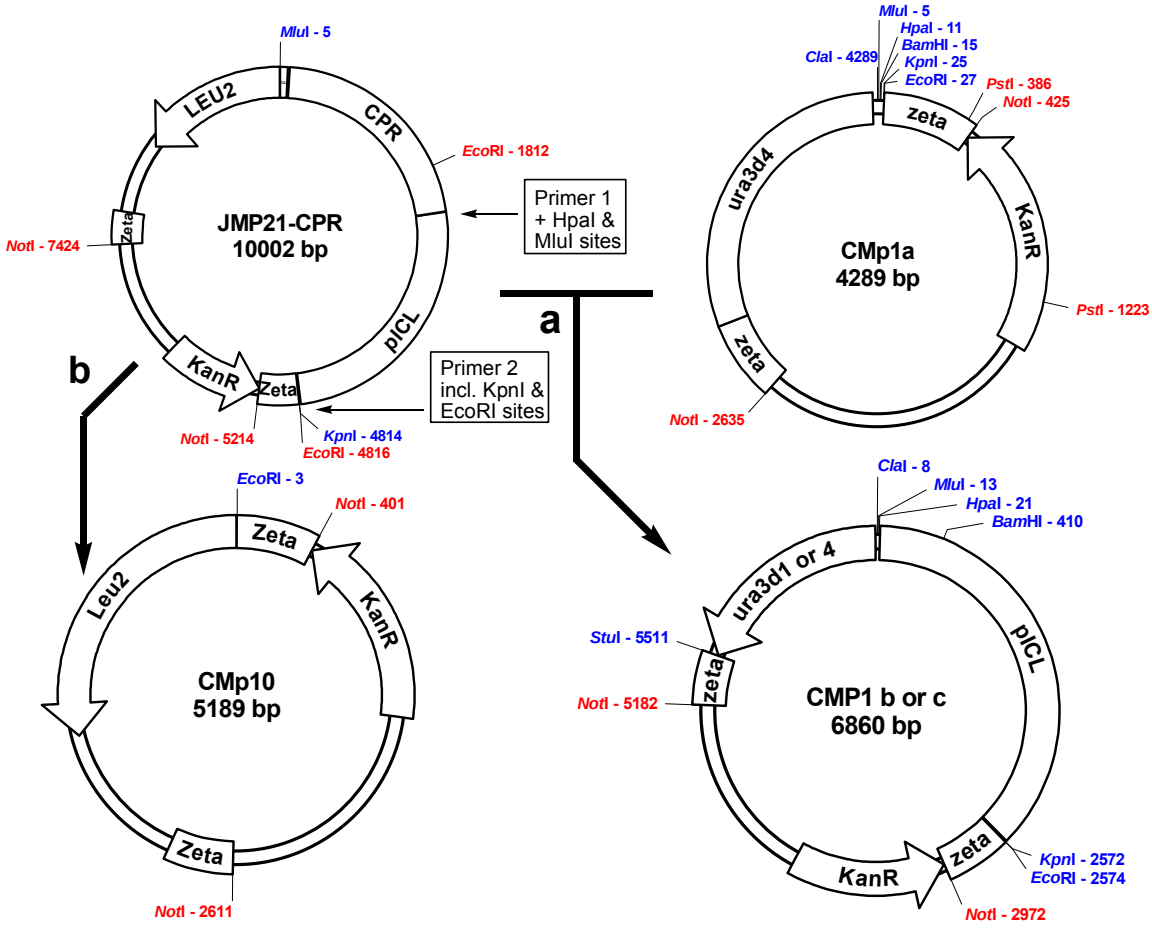


Figure 3.8: Construction of vectorsCMp1b, CMp1c and CMp10. (a) Insertion of the complete *ICL* promoter, amplified from JMp21-CPR, into JMp5 or CMp1a using *Mlu*I and *Kpn*I, forming CMp1b or CMp1c respectively. (b) Formation of CMp10 after removal of the *ICL-CPR* cassette from JMp21 using *Mlu*I and *Kpn*I, and subsequent re-ligation of the backbone.

The primers 53+LIP2tMluIF2 and 53+LIP2tClaIR2 (table 2.4.) were designed for the amplification of the *CYP53* gene along with the *LIP2t* terminator fragment from the JMp62-CYP53 vector. The PCR product included the 5' *Avr*II site preceding the start codon but excluded the *Eco*RI site at the 3' end of the *LIP2*t terminator. The primers allowed the introduction of a 5' *Mlu*I site and a 3' *Cla*I site. After sub-cloning in pGem-T Easy, the amplicon was

removed using *Mlu*I and *Cla*I and ligated into the CMp1b vector, producing vectors CMp3 (part a of figure 3.9.).

The *CYP53* gene in CMp3 is flanked by *Avr*II sites on either side, as it is in the JMp62 vector. These allowed the removal of *CYP53* and re-ligation of the backbone, producing the vector CMp4 (part b of figure 3.9). CMp4 contained the *ICL* promoter and the *LIP2*t terminator fragment on either side of a multiple cloning site for cloning other genes, including *Hpa*I, *Mlu*I and *Avr*II restriction sites.

The *CYP557* gene was amplified from the JMp62-*CYP557* vector using primers C557HpaIF and 557AvrIIR3 (table 2.4) which introduced an *Hpa*I restriction site at the 5' end and an *Avr*II restriction site at the 3' end. Following ligation of the PCR product into pGem-T Easy, this fragment was removed through restriction with *Hpa*I and *Avr*II and cloned into CMp4 using these enzymes, producing vector CMp5.

The corresponding multi-copy vectors CMp7-9 were produced similarly, starting with CMp1c.

The gene for *CYP102A1* was amplified from the pBM23 vector using the primers Bm3R1 and either BM3MluIF or Bm3F2BamHI (table 2.4), for the introduction of a 3' *Avr*II restriction site and either 5' *Mlu*I or BamHI restriction sites respectively. These restriction sites would be used for removal of the genes from the pGem-T Easy vector and cloning into further vectors. The *Mlu*I-*CYP102A1*-*Avr*II fragment was cloned into the CMp4 vector, producing CMp6 (Figure 3.9); while the *Bam*HI-*CYP102A1*-*Avr*II fragment was cloned into the JMp62 and JMp64 vectors, producing vectors CMp11 and CMp12 (Figure 3.10).

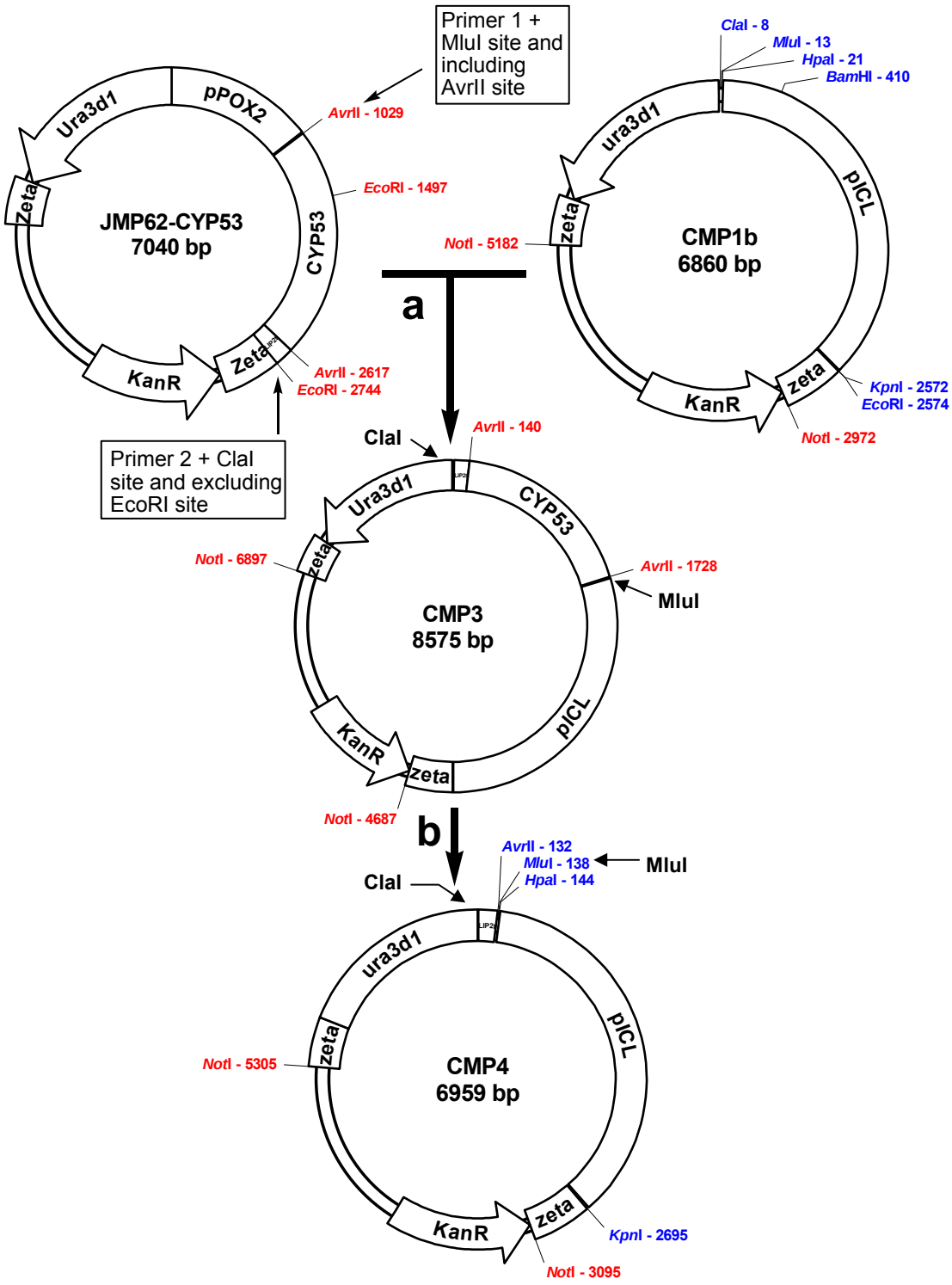


Figure 3.9: Cloning of the *CYP53-LIP2t* fragment, amplified from JMp62-CYP53, into CMp1b to form CMp3, and the subsequent removal of the *CYP53* gene from CMp3, producing CMp4 which can be used for further cloning. The *CYP53* gene (with its flanking *Avr*II sites) was amplified from the

JMp62-CYP53 vector along with the 3' LIP2t terminator region, with primers which introduced 5' *Mlu*I and 3' *Cla*I restriction sites. These restriction enzymes were used to incorporate the PCR product into CMp1b vector, fused to the *ICL* promoter, forming CMp3. The *CYP53* gene was also removed from this vector using *Avr*II, and the vector obtained by re-ligation of the backbone was termed CMp4.

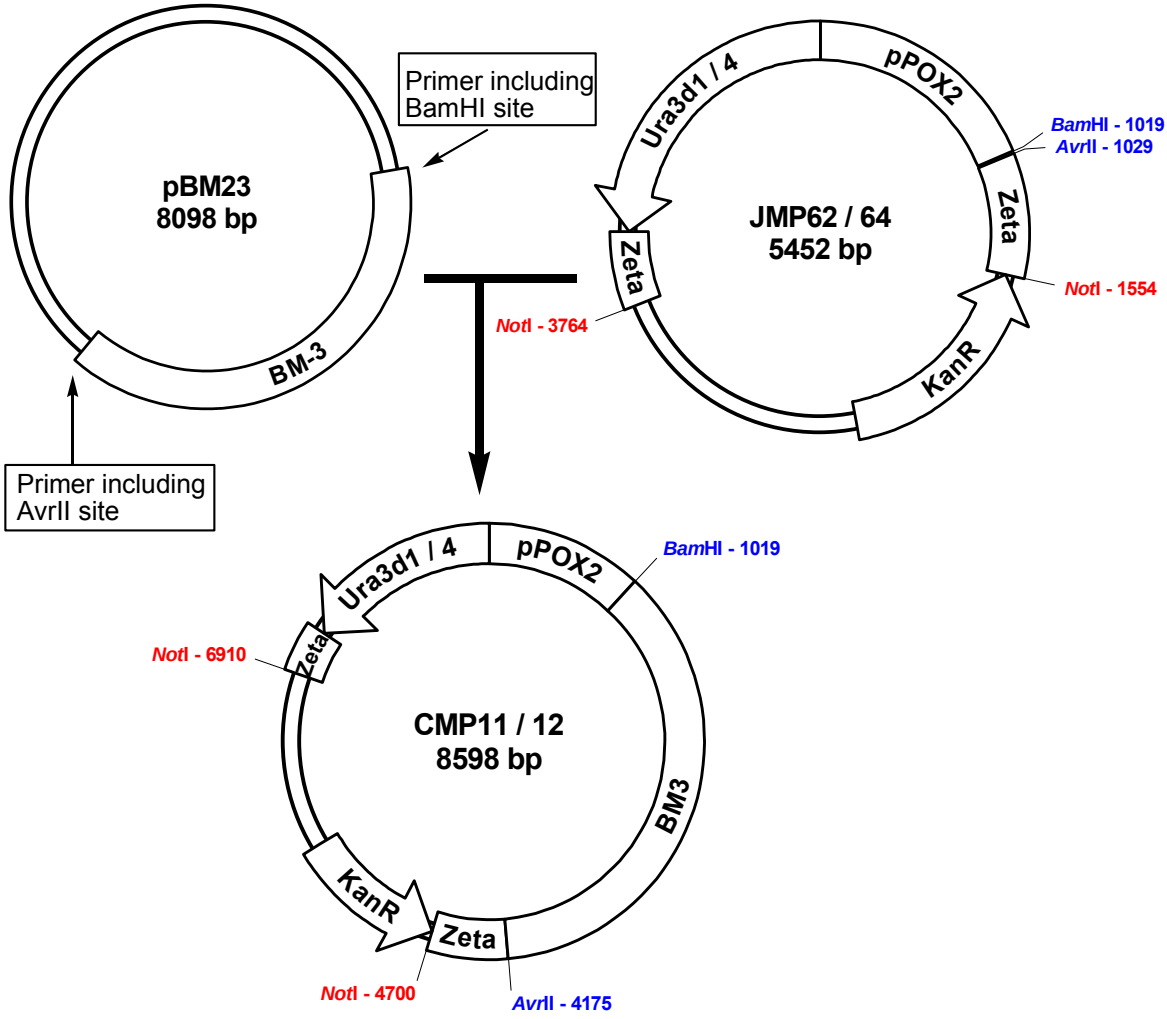


Figure 3.10: Construction of vectors CMp11 and CMp12 by ligation of CYP102A1 gene, amplified from pBM23, into JMp62 and JMp64. The *CYP102A1* gene was amplified from the pBM23 vector using primers which introduced 5' *Bam*HI and 3' *Avr*II restriction sites to the gene. These restriction enzymes were then used to fuse the *CYP102A1* gene to the *POX2* promoter

of vectors JMp62 and JMp64, producing vectors CMp11 and CMp12 respectively.

All of the cloning vectors used during this research contained *Not*I sites which flanked zeta elements in the vectors. These restriction sites allowed the linearization of the plasmid prior to transformation of *Y. lipolytica*, separating the bacterial portion of the vector from the yeast portion. The yeast section of the plasmid would then integrate homologously into corresponding zeta elements present in the genomes of some strains of *Y. Lipolytica*, or heterologously at a random point in genomes of strains devoid of zeta elements.

Table 3.1: Vectors constructed for integration and expression in *Yarrowia lipolytica*. All vectors contained zeta elements for integration into the host genome.

Vector	Backbone vector	Promoter	Cloned gene	Terminator	Marker
CMp2a	JMp5	*ICL* trunc[a]	None	None	*ura3d1*
CMp2b	CMp2a	*ICL* trunc[a]	*CYP53*	None	*ura3d1*
CMp2c	CMp2a	*ICL* trunc[a]	*CYP53*	*LIP2*t	*ura3d1*
CMp2d	CMp2a	*ICL* trunc[a]	*CYP53*	*LIP2*t	*ura3d1*
CMp2e	CMp2a	*ICL* trunc[a]	*CYP53*	None	*ura3d1*
CMp1a	JMp5	None	None	None	*ura3d4*
CMp1b	JMp5	*ICL*	None	None	*ura3d1*
CMp1c	CMp1a	*ICL*	None	None	*ura3d4*
CMp3	CMp1b	*ICL*	*CYP53*	*LIP2*t	*ura3d1*
CMp4	CMp3	*ICL*	None	*LIP2*t	*ura3d1*
CMp5	CMp4	*ICL*	*CYP557*	*LIP2*t	*ura3d1*
CMp6	CMp4	*ICL*	*CYP102A1*	*LIP2*t	*ura3d1*
CMp7	CMp1c	*ICL*	*CYP53*	*LIP2*t	*ura3d4*
CMp8	CMp7	*ICL*	None	*LIP2*t	*ura3d4*
CMp9	CMp8	*ICL*	*CYP557*	*LIP2*t	*ura3d4*
CMp10	JMp21-CPR	None	None	None	*Leu2*
CMp11	JMp62	*POX2*	*CYP102A1*	*LIP2*t	*ura3d1*
CMp12	JMp64	*POX2*	*CYP102A1*	*LIP2*t	*ura3d4*

[a] Truncated versions of the *ICL* promoter were used in these cases. More detail is given in the text (section 3.1.1.).

Sequencing of appropriate PCR fragments confirmed that the transition area between the promoters and genes in these vectors were confirmed to be in accordance with our vector maps.

3.2. Transformation of *Yarrowia lipolytica*

The *Y. lipolytica* strains that were used for the cloning of *CYP53B1, CYP557A1* and *CYP102A1* were TVN496 and FT120. TVN496 already contained an additional *CPR* gene under control of the *ICL* promoter and there were already derivatives of TVN496 available which containined *CYP557* under control of the *POX2* promoter. *Y. lipolytica* FT120 is a strain with β-oxidation completely blocked by disruption of all six fatty acyl-CoA oxidase encoding genes. This strain also has an additional *CPR* gene cloned, but under control of the *POX2* promoter. It thus also required the cloning of an additional *CPR* gene under control of the *ICL* promoter. The self-sufficient P450 encoded by CYP102A1 does not require an additional CPR and this gene could also be cloned into *Y. lipolytica* E150, the predecessor of TVN496. In addition, control strains also had to be created by transforming *Y. lipolytica* strains with parent vectors devoid of expression cassettes.

As is illustrated in figure 3.11, there are dramatic differences in transformation efficiencies when using zeta element containing vectors for transformation of *Y. lipolytica* strains containing zeta elements (e.g. strain E150 from the French inbreeding line) and those without zeta elements in their genomes (e.g. strain FT-120 derived from wild type strain W29 *via* strain Po1d).

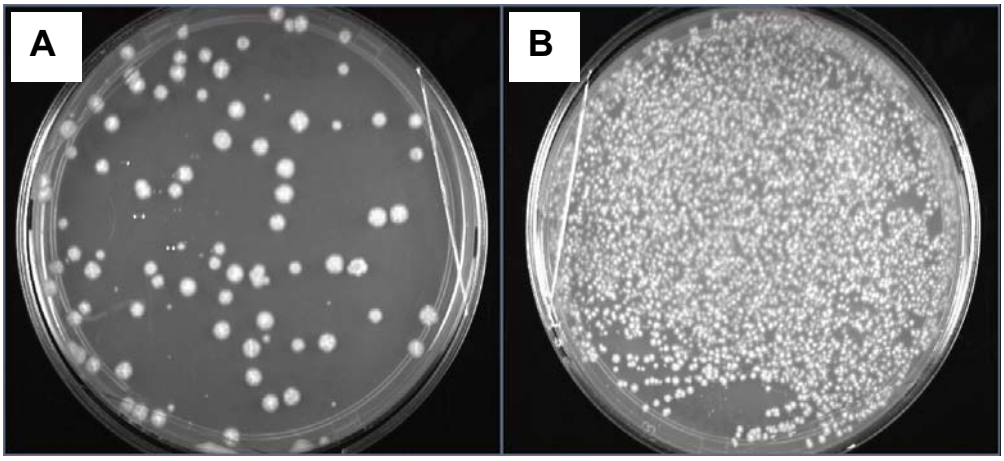


Figure 3.11: Differences in transformation efficiencies between *Yarrowia lipolytica* strains from different backgrounds Strains FT120 (a) and E150 (b) were transformed with the JMp62-CYP53 control vector.

This difference is due to the absence of zeta elements in the Po1d line, making genomic integration complicated and random, when it does occur. Furthermore, obtaining multiple copy integration, even in the strains possessing zeta elements, tends to be far more difficult than obtaining monocopy integrations (Juretzek *et al*; 2001). The FT-120 strain could not be omitted from this research though, as it had its entire β-oxidation pathway blocked, and hence held special interest to us for some applications.

Due to the transformation efficiency hurdles experienced, not all of the desired strains have been produced yet. The successful strains thus far, which have been confirmed through appropriate genomic DNA PCRs, are described in table 3.2.

Table 3.2: List of strains constructed in this study.

Strain	Host	Vector	Promoter: gene	Integration copy number	Auxotrophy
CTY001	E150	JMp5	n.a.	Single	leu$^-$
CTY002	E150	CMp10	n.a.	Single	ura$^-$
CTY003	CTY01	CMp10	n.a.	Single	None
CTY004	TVN496	JMp62	n.a.	Single	None
CTY005	TVN496	CMp4	*ICL*: no gene	Single	None
CTY006	TVN496	JMp62-CYP53	*POX2*: *CYP53*	Single	None
CTY007	TVN496	CMp3	*ICL*: *CYP53*	Single	None
CTY008	TVN496	CMp2b	*ICL*tr: *CYP53*	Single	None
CTY009	TVN496	CMp2c	*ICL*tr: *CYP53*	Single	None
CTY010	TVN496	CMp2d	*ICL*tr: *CYP53*	Single	None
CTY011	TVN496	CMp2e	*ICL*tr: *CYP53*	Single	None
CTY012	TVN496	CMp7	*ICL*: *CYP53*	Multiple	None
CTY013	TVN496	JMp62-CYP557	*POX2*: *CYP557*	Single	None
CTY014	TVN496	CMp5	*ICL*: *CYP557*	Single	None
CTY015	TVN496	JMp64-*CYP557*	*POX2*: *CYP557*	Multiple	None
CTY017	TVN496	CMp9	*ICL*: *CYP557*	Multiple	None
CTY018	TVN496	CMp11	*POX2*: *CYP102A1*	Single	None
CTY019	TVN496	CMp6	*ICL*: *CYP102A1*	Single	None
CTY020	TVN496	CMp12	*POX2* : *CYP102A1*	Multiple	None
CTY021	FT-120	JMp21-CPR	*ICL*: *CPR*	Single	ura$^-$
CTY022	FT-120	CMp10	n.a.	Single	ura$^-$
CTY023	FT-120	JMp62-*CYP53*	*POX2* : *CYP53*	Single	leu$^-$
CTY024	FT-120	JMp62-*CYP557*	*POX2* : *CYP557*	Single	leu$^-$
CTY025	FT-120	CMp11	*POX2* : *CYP102A1*	Single	leu$^-$
CTY026	CTY021	JMp5	n.a.	Single	None
CTY027	CTY021	CMp4	*ICL*: no gene	Single	None
CTY028	CTY021	CMp5	*ICL*: *CYP557*	Single	None
CTY029	CTY022	CMp4	*ICL*: no gene	Single	None
CTY030	CTY022	CMp6	*ICL*: *CYP102A1*	Single	None

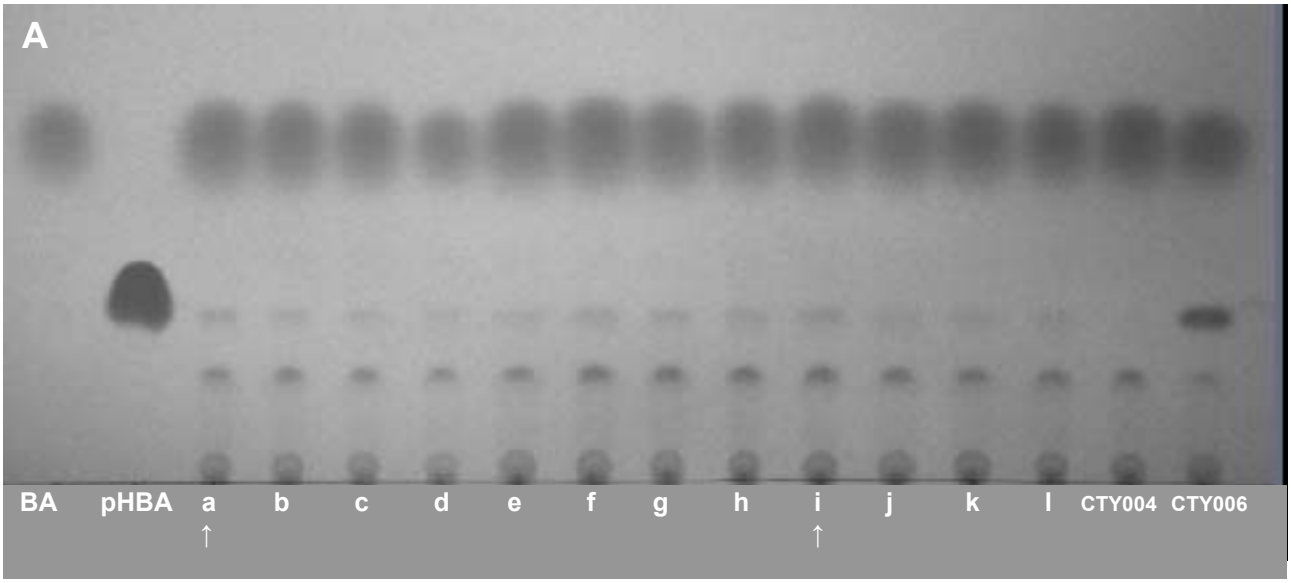
In table 3.2., 'n.a.' indicates strains transformed with vectors lacking an expression cassette, but containing a marker gene for prototrophy restoration. Such strains served as either negative control strains during activity assays, or as intermediate strains awaiting further transformation using another vector.

Biotransformations using recombinant strains

Recombinant strains of *Y. lipolytica* had been prepared containing genes for *CYP53* as a reporter gene, *CYP557* for which the product's activity was yet to be exactly determined, and the self-sufficient bacterial P450 *CYP102A1.* With these strains available, activity assays would be performed to determine the modified activities of these strains, as compared to control strains without heterologous P450s expressed. Since the focus in our research group is largely devoted to whole-cell systems, whole cell biotransformations were predominantly focussed on.

3.3. Confirmation of benzoate-*para*-hydroxylase activity of strains which had *CYP53* cloned

Transformants obtained from the transformation of strains TVN496 and FT120 with vector CMp3 (which contained the CYP53 - LIP2t PCR product fused to the complete *ICL* promoter) were screened for benzoate-*para*-hydroxylase activity. The biotransformations were done in YPD medium, with ethanol (2% v/v) used for induction of the *ICL* promoter in the test strains, and oleic acid (2% v/v) as inducer for control strains. Inducers were added only once, after the first 24 hours of main culture growth. Benzoic acid (0.5% w/v) was added after 24 hours of induction. Figure 3.12 and Fig. 3.13 shows the TLC analysis of the extracts from these initial screens of TVN496 and FT120 transformants respectively.



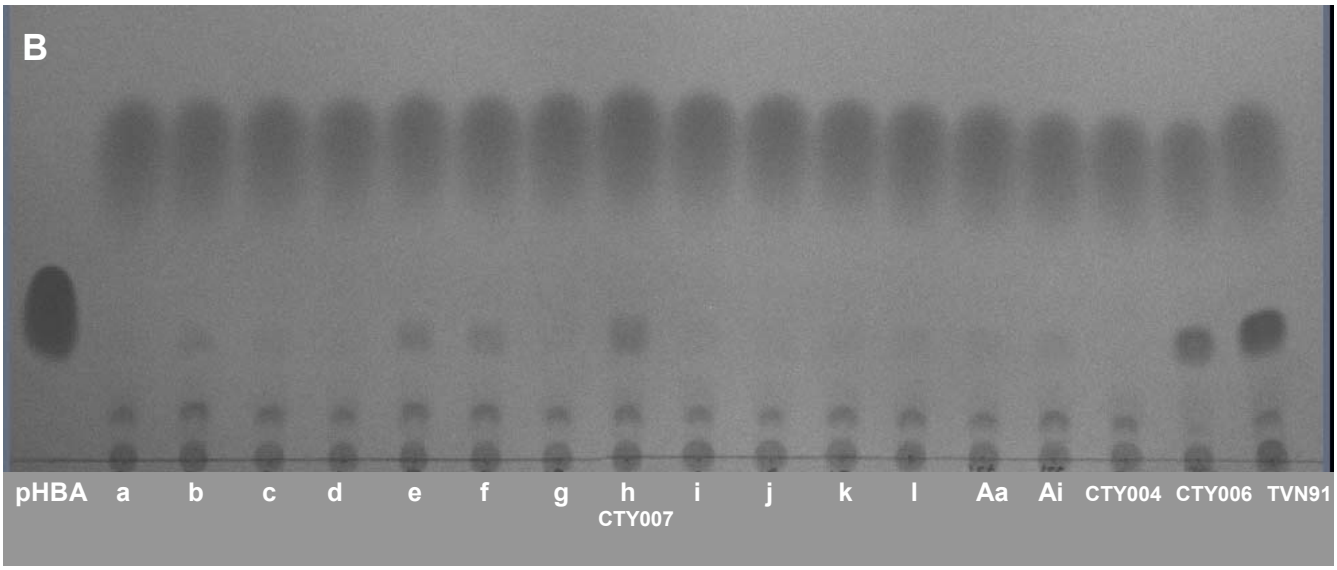


Figure 3.12: TLC analysis of benzoate-*para*-hydroxylase activities in *Y. lipolytica* TVN496 strains transformed with CYP53 under control of the complete ICL promoter. BA – benzoic acid, pHBA – *para*-hydroxy benzoic acid. Plate A - 72h extracts from first twelve transformants containing *ICL*:*CYP53*, CTY004 (negative control) and CTY006 (positive control containing *POX2*:*CYP53*). Plate B - 48h extracts from another twelve transformants

containing *ICL*:*CYP53*, Aa and Ai best strains from previous screen, CTY004, CTY006 and TVN91.

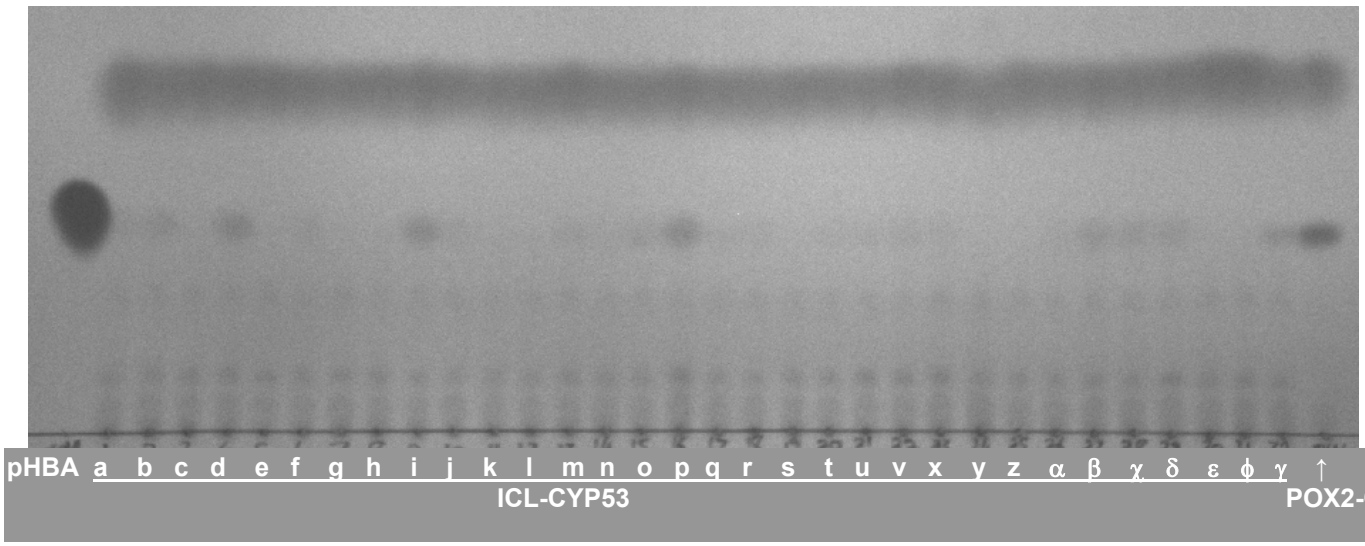


Figure 3.13: TLC analysis of initial benzoate-para-hydroxylase activity in *Y. lipolytica* strain FT-120 transformed with CYP53 under control of the complete ICL promoter. pHBA – *para*-hydroxy benzoic acid; 48h extracts from thirty transformants (a - γ) containing *ICL:CYP53* and one transformant containing *POX2:CYP53* (↑).

Strain CTY006 with CYP53 under control of the *POX2* promoter (transformed with JMp62-CYP53) still performed better than any of the *ICL:CYP53* transformants. One transformant with *CYP53* under the *ICL* promoter (Fig. 3.12 Plate B, sample h) showed far higher activity, even at 48 hours after substrate addition, than the other TVN 496 transformants. This strain, labelled CTY007, was promising enough to be compared with CTY006, in further biotransformation experiments.

A number of FT120 transformants (i.e. samples d, i and p in Fig 3.13) containing *ICL*:*CYP53* showed, after induction with ethanol, activity that was almost comparable to the activity observed with the strain containing

POX2:CYP53 induced with oleic acid. It should be noted that these strains did not contain a *CPR* gene under control of the *ICL* promoter.

Only two transformants were isolated from transformations of TVN496 with the multiple-copy JMp5-ICL-CYP53 vector. Their genomic DNA was isolated and checked for the presence of the gene using PCR (figure 3.14.).

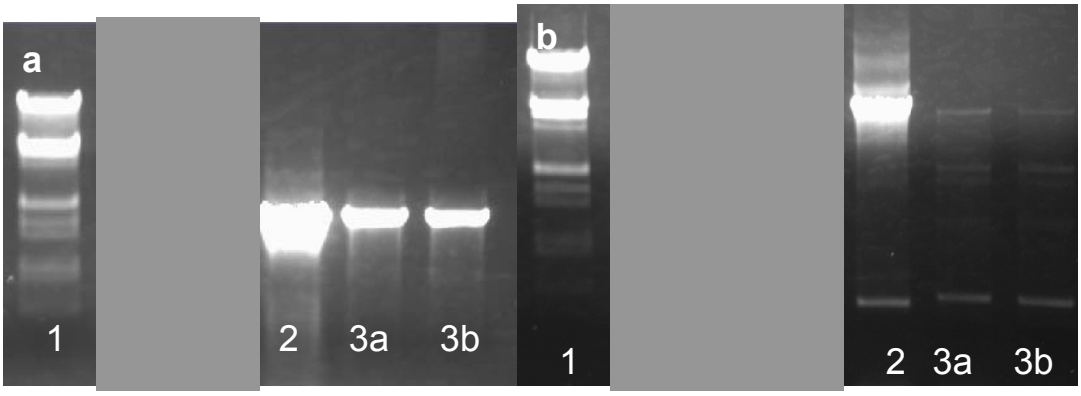


Figure 3.14: PCR reactions on potential multi copy *ICL:CYP53* transformants. (A) Amplification of only the CYP53 gene. (B) Amplification of the complete *ICL:CYP53* fragment. Lane 1 λIII marker; lane 2 vector DNA control; lanes 3a & b the two tested isolates.

The PCR results were convincing enough for us to proceed with screening of these strains. They were screened in the same way as the single copy transformants, and a TLC plate is shown in figure 3.15. The activities obtained seemed very low in comparison to the previous results obtained. Nevertheless, one of these multi-copy transformants, labelled CTY012, was selected for comparison with other strains.

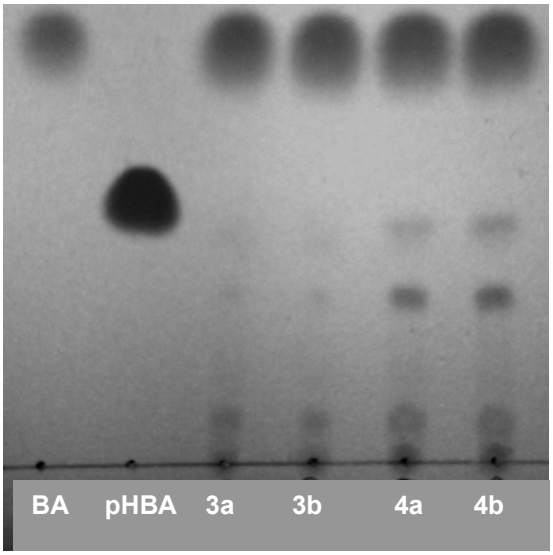


Figure 3.15: TLC analysis for benzoate-*para*-hydroxylase activity in multi-copy transformants of TVN496 with *ICL:CYP53*. BA – benzoic acid, pHBA – *para*-hydroxy benzoic acid. Samples 3a and 3b were taken directly after substrate addition, and Sample 4a and 4b 72 hours after substrate addition.

Strains CTY007 and CTY012, *ICL:CYP53* single and multi-copy transformants respectively from TVN496 and CTY006 a *POX2:CYP53* transformant from TVN496 were selected for further comparisons with TVN91, a *POX2:CYP53* multi-copy transformant from previous studies, and strain CTY004, a negative control strain obtained by transformation of strain TVN496 with JMp62.

The strains were prepared for biotransformations in YPD medium as described before. After 24 hours, ethanol (2% w/v) and oleic acid (2% v/v) were added as inducers to separate flasks of each strain. Benzoate (0.5 %) was added to all flasks after 6 h. After 18 hours a further 1% oleic acid was added to all flasks, as well as 1% ethanol to the flasks which were initially induced with ethanol. Thereafter a further 0.5% of only oleic acid or ethanol was added to the appropriate flasks at 24 h intervals.
Samples were taken at 24 hour intervals from the time of substrate addition. The results of these comparative biotransformations are shown in figure 3.16.

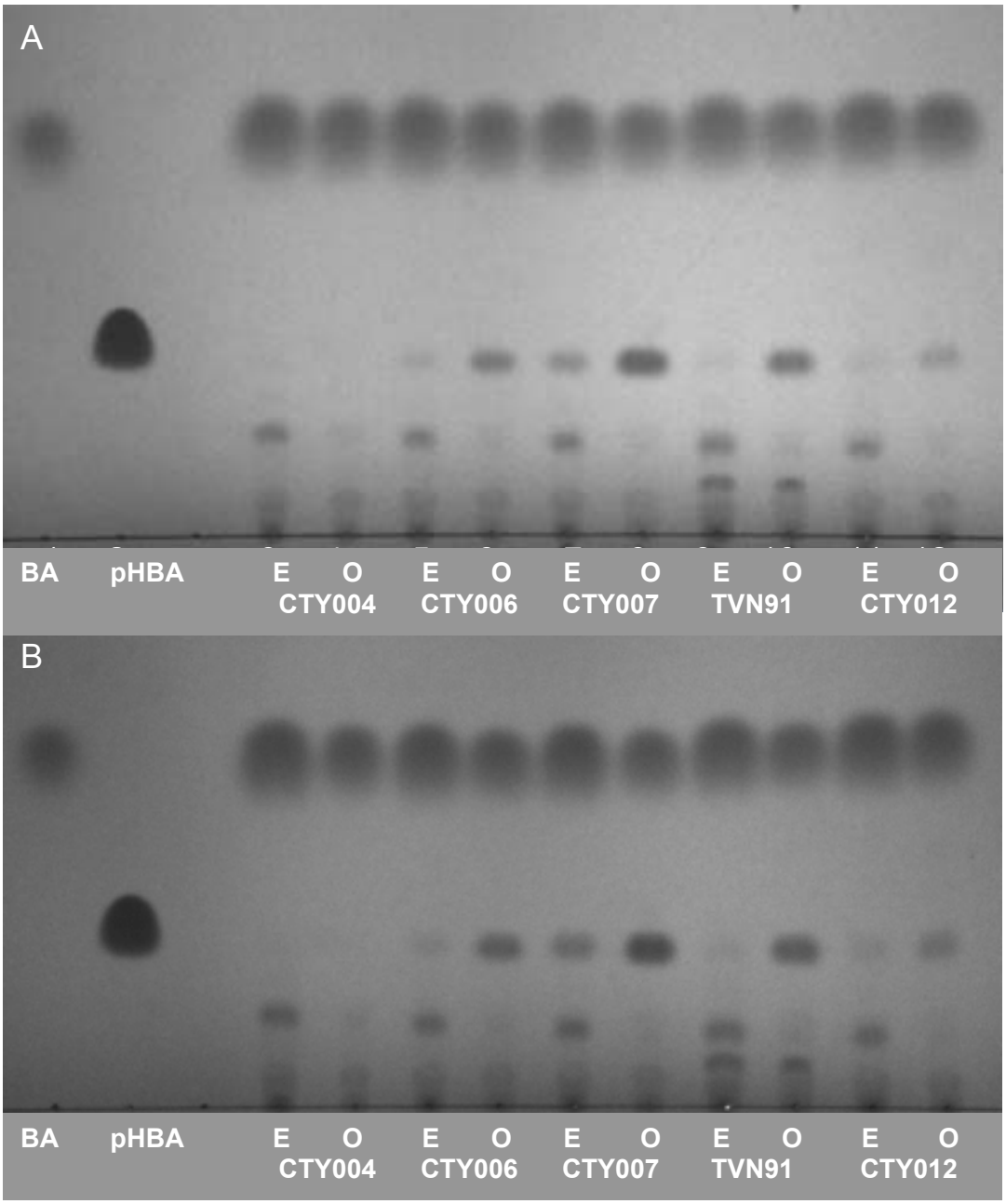


Figure 3.16: TLC plates for the biotransformation of benzoic acid by various strains. Plate A 78 h extracts and plate B 100 h extracts of biotransformations done with Y. *lipolytica* strains CTY004, CTY006, CTY007, TVN91 and CTY012 after induction with mainly ethanol (E) or oleic acid (O). BA – benzoic acid, pHBA – *para*-hydroxy benzoic acid.

Firstly, the negative control strain CTY004 showed no benzoate-*para*-hydroxylase activity as expected. In the results for the mono-copy strains CTY006 and CTY007 it is observed that when the *POX2* controlled gene is induced by oleic acid, and the *ICL* controlled gene is induced by ethanol, the results are as we obtained during the screening, i.e. that the strain with *CYP53* under *POX2* is performing better. However, when we compare both strains induced by oleic acid, the activity of the strain with *CYP53* driven by the *ICL* promoter surpasses that of the strain with *CYP53* driven by the *POX2* promoter.

The strain CTY012 containing multiple copies of *CYP53* driven by the *ICL* promoter is still performing poorly, when compared to the multi copy strain TVN91, which contains multiple copies of *CYP53* driven by the *POX2* promoter. Strain CTY012 is even performing worse than the single copy transformants. The strain has either not reached stability with its copy number (Juretzek *et al*, 2001), or does not have the high copy number expected. Nevertheless, it showed better activity when induced with oleic acid than with ethanol.

Ethanol does not induce the *POX2* promoter (Juretzek *et al*, 2000). Probably the only reason we saw any activity at all with the *POX2:CYP53* strains induced with ethanol, is due to the once-off addition of oleic acid to these cultures. This was, in retrospect, a flaw in the experimental design. One thing which is clear is that the *ICL* controlled genes are performing better after induction with oleic acid than ethanol. In the work done by Nthangeni *et al* they only used fatty acids to induce both promoters, and found similar activities of their genes expressed under either promoter (Nthangeni *et al*, 2004).

While it is not at all surprising that the *ICL* promoter is induced by oleic acid, Juretzek *et al* found similar activities after induction with ethanol and oleic acid, not necessarily better activity with either (Juretzek *et al*, 2000). That information,

together with our improved activity under *ICL* using oleic acid instead of ethanol, indicates that the differences found in this study may not simply be due to one promoter being more effective than the other. It becomes more likely that the problems encountered in our study are more indirect, possibly involving ethanol repression at some other point in the pathways of substrate degradation. When fatty acids are used, all the relevant pathways necessary for the metabolism of the usual substrates are activated. This is unlikely to be the case when ethanol is used. Possibly one or more of the transporters involved in metabolite relocation within the cell are affected negatively by the ethanol, or just not switched on by this substance. Then the activity differences would be related to inferior uptake of substrates to begin with, or diminished passage of metabolites within the cell (Thevenieau *et al*; 2006).

3.4. Activity analysis of strains in which *CYP557A1* was cloned, for modified hydroxylase activity

TVN 496 derivatives

Genomic DNA was isolated from TVN496 transformants after transformation with CMp5, and appropriate PCRs were performed to test for the presence of the cloned gene (figure 3.17).

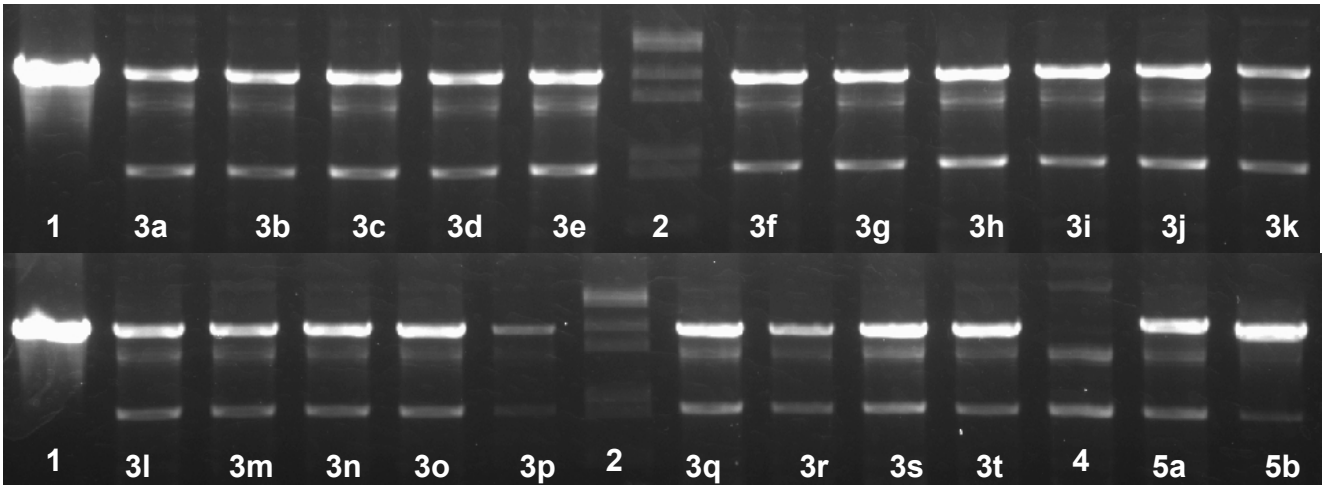


Figure 3.17: PCR results for *CYP557A1* amplification from genomic DNA of TVN496 transformants containing *CYP557A1*. Lanes 1 vector control; lanes 2 λIII gene marker; lanes 3a – 3t TVN496 transformants containing *ICL:CYP557A1*; lane 4 CTY004 (negative control); lane 5a TVN210 and lane 5 b TVN348 (positive controls containing *POX2:CYP557A1*).

Since all of the tested transformants gave the expected PCR products of approximately 1.6 kbp, they were all screened for activity towards hexylbenzene. It has been shown that alkylbenzenes are useful substrates for rapid screening using TLC analysis, since the substrates and products are UV active (van Rooyen 2005; Obiero 2006).

The *ICL:CYP557A1* transformants and the negative control strain CTY004 were induced for 24 hours with ethanol (2% v/v), while the TVN210 strain was induced with only oleic acid (2% v/v) and the TVN348 strain with oleic acid and ethanol simultaneously (2% v/v each). After 24 h of induction substrate (1% v/v hexylbenzene) was added. Ethanol was not added again during this screen. Samples were taken regularly for further analysis.

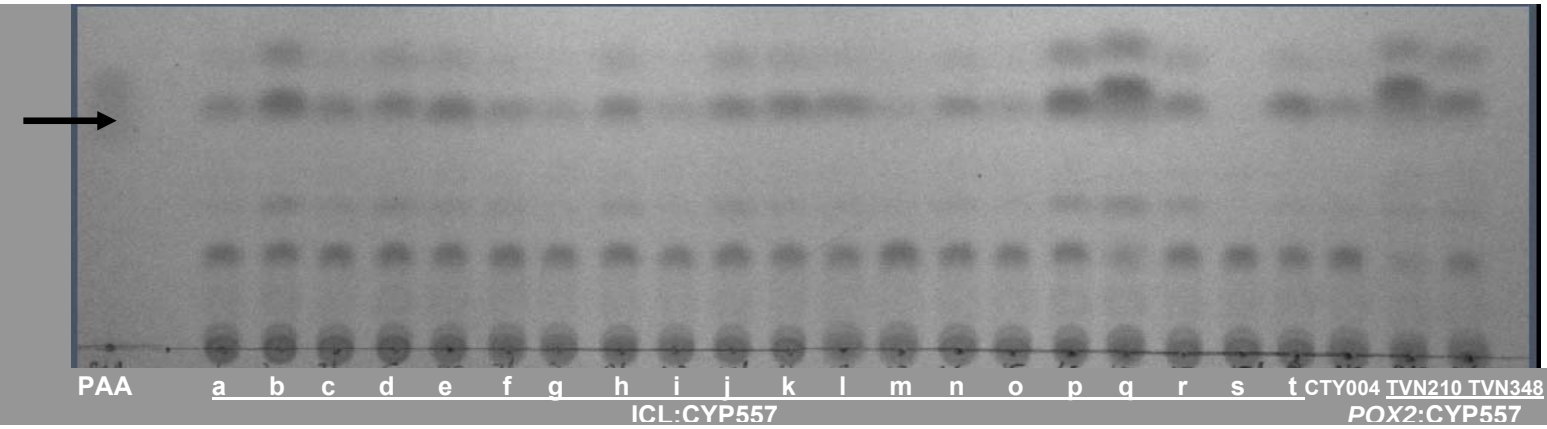


Figure 3.18: TLC plate of initial screening of TVN 496 transformants containing *CYP557A1*. PAA - phenylacetic acid standard; Test samples are 24h extracts of TVN496 transformants containing *ICL:CYP557A1* (twenty strains a - t), CTY004 (negative control transformed with JMp62); TVN210 (mono copy *POX2:CYP557A1*) and TVN348 (multi-copy *POX2:CYP557A1*).

The results (figure 3.18) show at least three promising *ICL:CYP557A1* transformants (strains b, p and especially q). GC analysis was performed on samples from the biotransformations with these 3 strains as well as the control strains (figure 3.19). GC analysis confirmed that strain q was throughout the experiment the best converter of hexylbenzene to phenylacetic acid. Transformants p and q also initially accumulated the intermediate phenylbutanoic acid. It was decided to proceed with further comparative biotransformations using strain q, renamed CTY014.

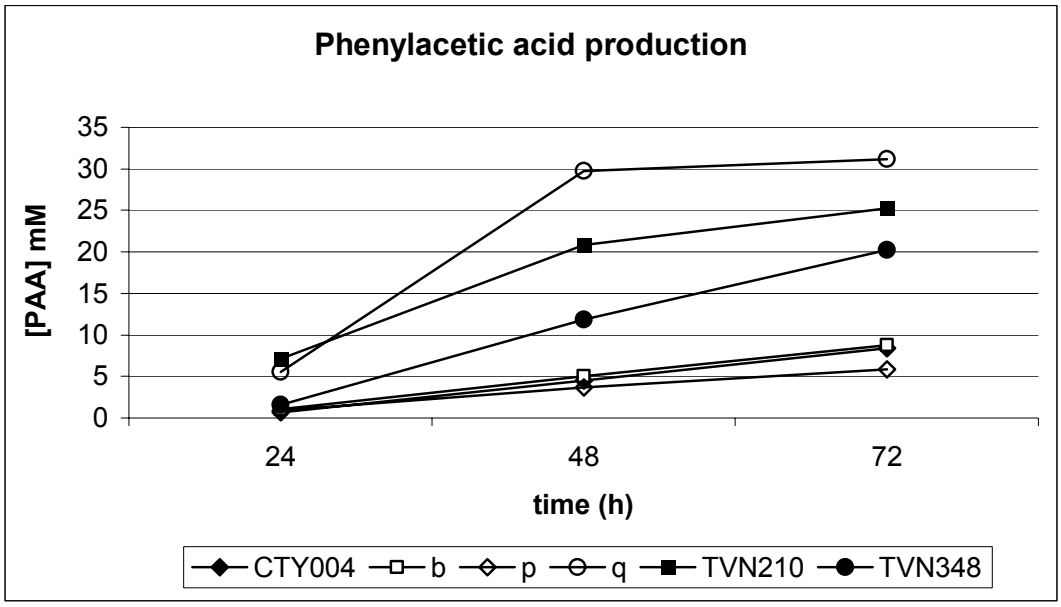


Figure 3.19: GC data obtained from hexylbenzene screen of *CYP557A1* transformants b (empty squares), p (empty diamonds), and q (empty circles); as well as the negative control strain CTY004 (solid diamonds), and strains TVN210 (solid squares), and TVN348 (solid circles), both containing *POX2:CYP557A1*, in monocopy and multi-copy, respectively.

The strain CTY013, a TVN496 strain transformed with JMp62-CYP557 in this study, along with the strain CTY014 and the negative control strain CTY004 were further compared for biotransformation of hexylbenzene. Ethanol and oleic acid (2% v/v) were both added as inducers to separate 24 h old cultures. After 6 h of induction hexylbenzene (HB) (1% v/v) was added as the substrate. After a further 18 h inducers were supplemented (1% v/v) and again after another 24 h (0.5% v/v). During the process, samples were collected regularly for GC analysis (figure 3.20).

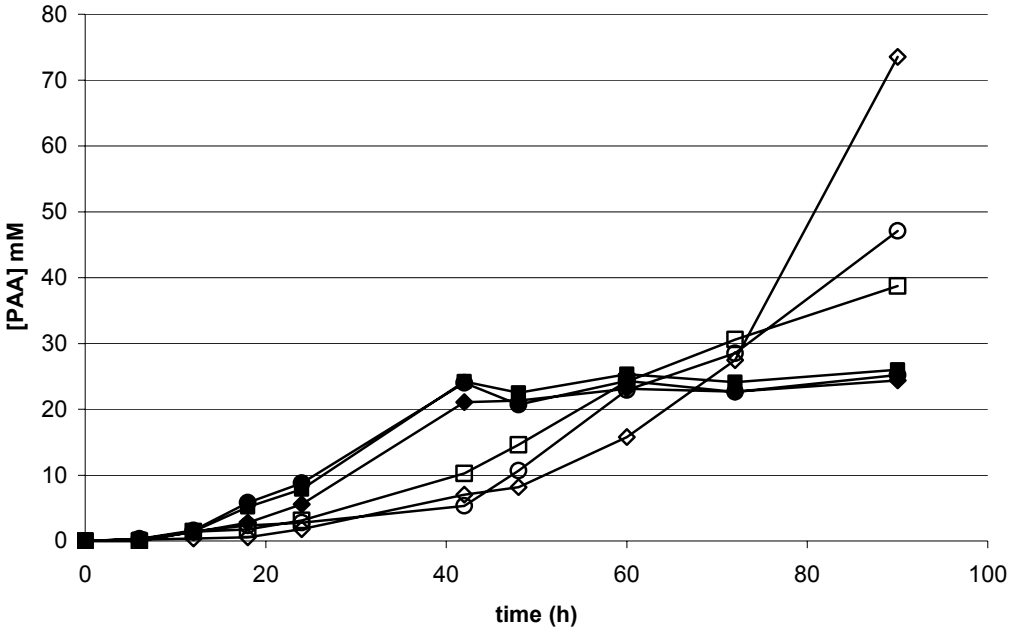


Figure 3.20: GC analysis of hexylbenzene bioconversion to phenylacetic acid by CTY004 (negative control) (circles), CTY013 (*POX2:CYP557*) (squares) and CTY014 (*ICL:CYP557*) (diamonds). Inducers were ethanol open symbols) and oleic acid (closed symbols) added after 24 h (2 % v/v); 48 h (1 % v/v) and 72 h (0.5 % v/v) growth. Hexylbenzene (1 % v/v; *ca.* 62mM) was added after 30 h growth, which is 0 h on the graph.

Substrate was added earlier than in the previous experiment, with the expectation that it will enhance the difference between wild-type and cloned P450 activity. However, the strains induced by oleic acid (closed symbols) followed a very similar pattern in phenylacetic acid production, regardless of gene cloned or promoter used. This is a strong indication that the observed activities in these cases are mainly wild-type activities.

The cases in which ethanol was the inducer (open symbols) were more complex though. When the cultures reached the point of about 42 hours after substrate addition, the ethanol concentration was expected to be rather low. Only 0.5% v/v

additional ethanol was re-added at that point. It can be observed that a steady increase in phenylacetic acid (PAA) production occurred around that point, indicating that the repression of the native P450 enzymes by ethanol was wearing off. The negative control even overtook the *POX2:CYP557* containing CTY013 strain in terms of PAA production. Initial excitement over the rapid increase in activity in the *ICL:CYP557* containing CTY014 strain subsided when the product concentration exceeded the initial substrate concentration, signifying that something other than HB conversion to PAA was taking place in the culture after prolonged biotransformation times.

While the amount of accumulated product contributed by the cloned P450s is questionable, the delay in product formation was common for all the ethanol induced cells. The formation of the major product (PAA) only starts competing with the levels produced by the oleic acid induced cells after the ethanol concentration declines.

As was mentioned earlier, the results could be explained as the wild-type P450 inhibition by ethanol wearing off, but they could also be in accordance the results obtained for the CYP53 transformants. The presence of ethanol seems to impart a negative influence on the uptake and/ or transport of hydrophobic substances. Once the ethanol is consumed, the hexylbenzene enters the regular degradation pathway and possibly itself induces the P450s required for its initial hydroxylation. Perhaps degradation intermediates of this substrate also induce the ICL promoter, causing a double effect of cloned and native P450s reacting on the substrate.

Recent studies in our research group have shown that similar continual ethanol addition but at higher concentrations each time drastically lowered the product formation, regardless of the strain used. It remains unclear at this stage whether the effect was due to native P450 inhibition, substrate uptake and/ or transport inhibition, or due to toxicity effects.

Further studies in our group were carried out with only initial induction with ethanol (2% v/v) and no further supplementation with ethanol or any other inducer. Hexylbenzene and butylbenzene were used as substrates, and a clearer difference was observed using butylbenzene (Ramorobi, personal communication). Seemingly hexylbenzene, or one of its intermediate products, is capable of inducing the wild-type P450s, but butylbenzene seems to have no such induction capability of the wild-type P450s. The possibility that butylbenzene is a preferred substrate to hexylbenzene for CYP557A1 cannot be ruled out at this stage either.

Aside from the likely confirmation of the substrate uptake/ transport system inhibition found in the CYP53 tests, these CYP557 tests also brought to light another potential limitation, this time concerning the *POX2* promoter. Since production by ethanol induced strains eventually exceeded that of oleic acid induced strains, competition of inducers and substrates for the β-oxidation pathway seems to occur in cases where both require this pathway for degradation.

FT-120 derivatives

The FT-120 strain was transformed with the vector JMp21-CPR resulting in the strain CTY021. This was done to complement the leucine auxotrophy in the FT-120 strain, and add an additional copy of the *CPR* gene under control of the *ICL* promoter. The CTY021 strain was then further transformed with CMp5. Only three transformants could be rescued from the transformation plate, and they were subjected to genomic DNA PCR (figure 3.21).

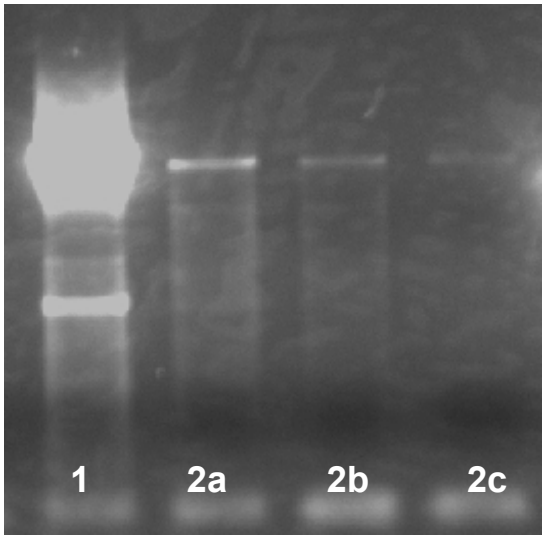


Figure 3.21: PCR results for CYP557 amplification from genomic DNA of CTY021 transformants containing *ICL*:*CYP557*. Lane 1 vector control; lanes 2a – 2c CTY021 transformants containing *ICL*:*CYP557.*

Since the FT-120 parent strain has its β-oxidation completely blocked, the transformants were tested for dioic acid (DCA) production from dodecane (figure 3.22). The strains were induced once with ethanol (2% v/v) after 24h growth and substrate was added 24h later.

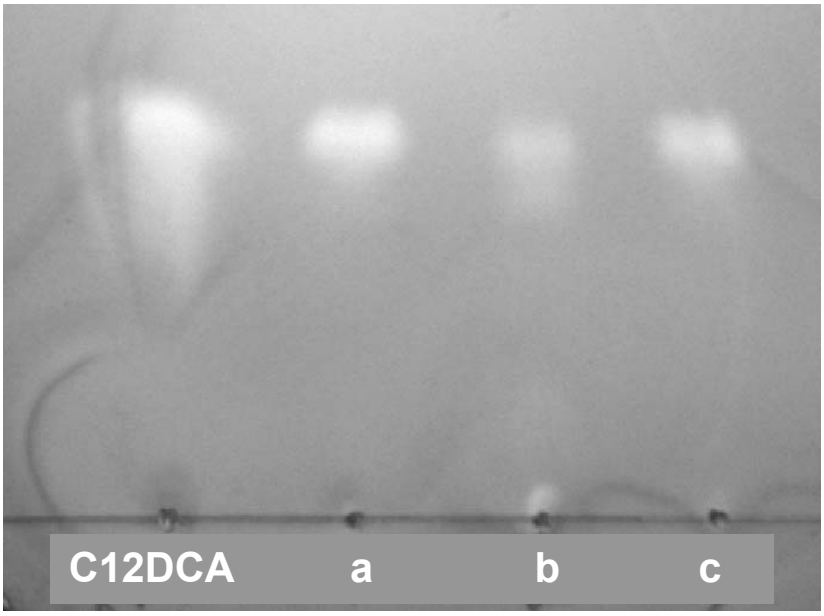


Figure 3.22: TLC analysis for conversion of dodecane to dodecane dioic (C12DCA) by CTY021 transformants containing *ICL*:*CYP557*. Test samples (a – c) are extracts from samples taken 72 h after addition of dodecane (1% v/v)

At 18 h after substrate addition strain b showed some promise. The observed "product" however, was also present in the sample taken immediately after substrate addition, and did not seem to increase as time progressed. At 72 h, (figure 3.22) the dioic acid production of strains a and c seemed to have surpassed that of b. GC analysis confirmed that this "product" of b was indeed not dodecane dioic acid (results not shown). Transformant a was taken instead for future use, and renamed CTY028.

FT-120 was also transformed with CMp10, still complementing the auxotrophic marker but without adding an extra CPR copy. This strain was labelled CTY022. CTY021 and CTY022 were both transformed with CMp4, giving strains CTY027 and CTY029 as negative controls for CTY028. These 3 strains were then compared for dioic acid accumulation from dodecane.

Either glucose or ethanol (0.5% v/v) was added together with the substrate dodecane (1% v/v) to 24 h cultures, and was supplemented again to the same

final concentrations every 24 hours thereafter. The substrate was only added once. The pH was maintained in each case at pH 8. Samples were taken regularly for analysis (figure 3.23).

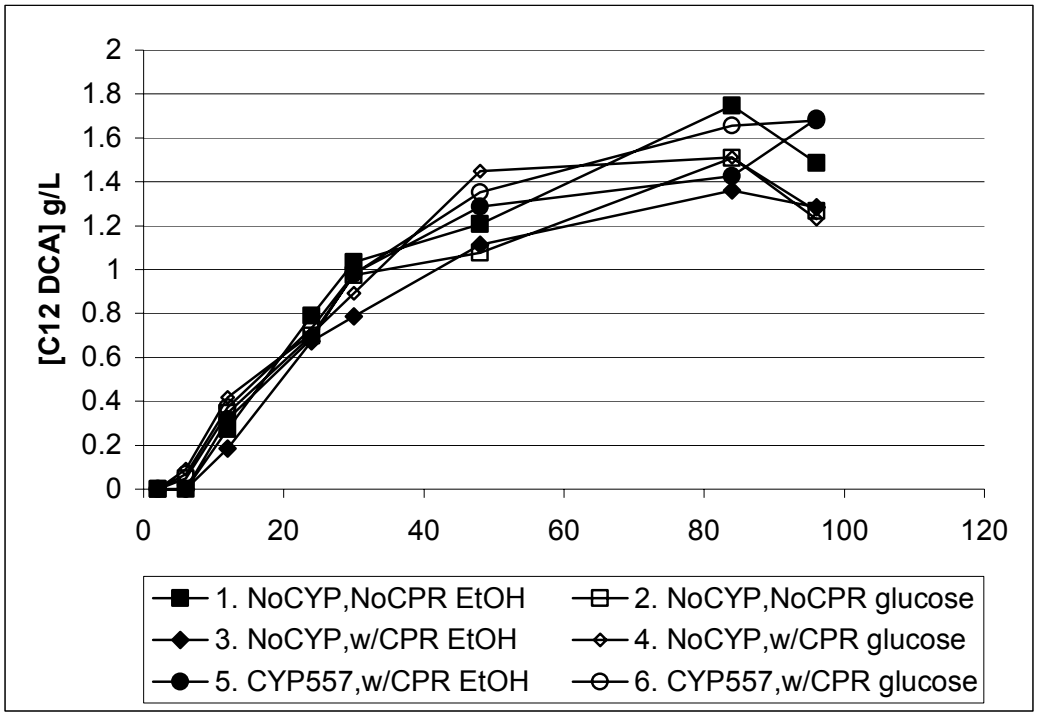


Figure 3.23: GC analysis of dodecane bioconversion to dodecane dioic acid by CTY029 (negative control, no additional CPR or CYP557 cloned) (squares), CTY027 (negative control, *ICL:CPR,* no additional CYP557 cloned) (diamonds) and CTY028 (*ICL:CPR* and *ICL:CYP557*) (circles). Ethanol (open symbols) and glucose (closed symbols) were added (0.5 % v/v) after 24 h; 48 h and 72 h growth. Dodecane (1 % v/v) was added after 24 h growth, which is 0 h on the graph.

Dioic acid production in this experiment was disappointingly low and no clear differences could be observed between the different strains. Dodecane dioic acid production of up to 8 g/l has already been reported with *Y. lipolytica* MTLY37, a strain that is only partially blocked in β-oxidation (Smit *et al*, 2005). In further optimization studies in our research group dodecane dioic acid

concentrations of up to 14 g/l has been achieved with CTY029 in a bioreactor (Gumede, personal communication). Once requirements for dioic acid production is better understood and has been optimized the effect of the additional CPR and CYP557 genes under the ICL promoter will again be evaluated. Further optimization studies of this process are under way in our research group.

3.5. Analysis of activity of strains in which *CYP102A1* was cloned

TVN496 derivatives

Integration of *POX2:CYP102A1* and *ICL:CYP102A1* into TVN496 was confirmed by PCR amplification of an expected 3.15 kbp product from genomic DNA of transformants (Fig. 3.24)

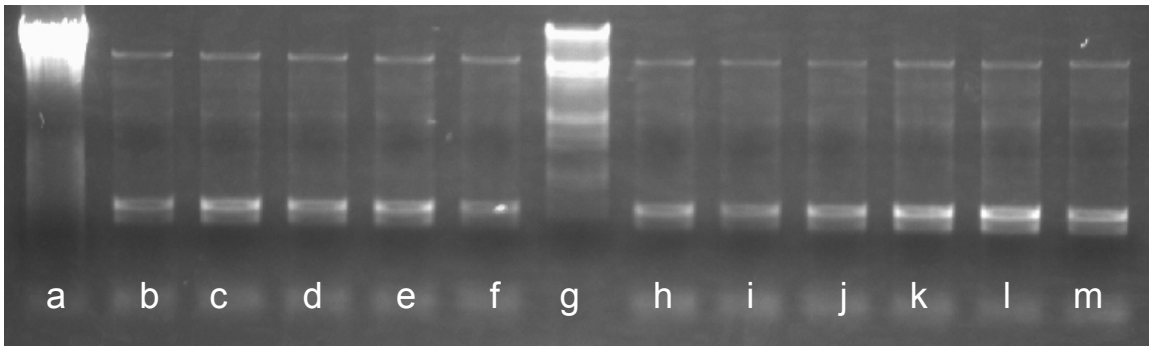


Figure 3.24: PCR Genomic DNA PCR results for TVN496 strains transformed with *CYP102A1* under the *POX2* or *ICL* promoters. Lane a vector control, lanes b-f and h-m TVN496 transformants containing *CYP102A1*, lane g the λIII gene marker.

Testing subterminal hydroxylation of fatty acids with TVN496 derivatives which has in tact β-oxidation was not feasible. A commonly used assay for screening of CYP102A1 activity involves *p*-nitrophenoxycarboxylic acids (pNCA) as substrates (Schwaneberg *et al*, 1999). The product formed when using

CYP102A1 is *p*-nitrophenolate, which is easily detected using a spectrophotometer at 410 nm (figure 3.25).

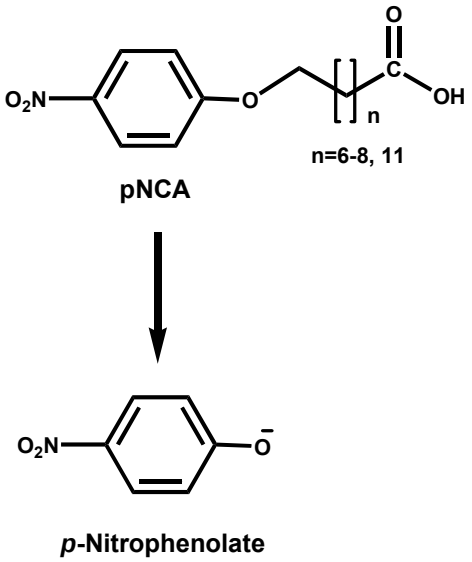


Figure 3.25: Conversion of *p*-Nitrophenoxycarboxcylic acids to *p*-Nitrophenol by CYP102A1.

pNCA is not commercially available. 4-Nonyloxybenzoic acid (NOBA) is commercially available and resembles pNCAs except that the functional group in the *para* position to the ether is a carboxyl group rather than a nitro group; and that the ether linked substituent is an alkane chain rather than a fatty acid (Fig. 3.26). It has been shown that Y. lipolytica strains with functional β-oxidation degrade NOBA via fatty acid intermediates (which more closely resemble pNCAs) *p*-hydroxybenzoic acid (pHBA) (Obiero, 2006). The conversion of NOBA to pHBA via its detectable intermediates is illustrated in figure 3.26.

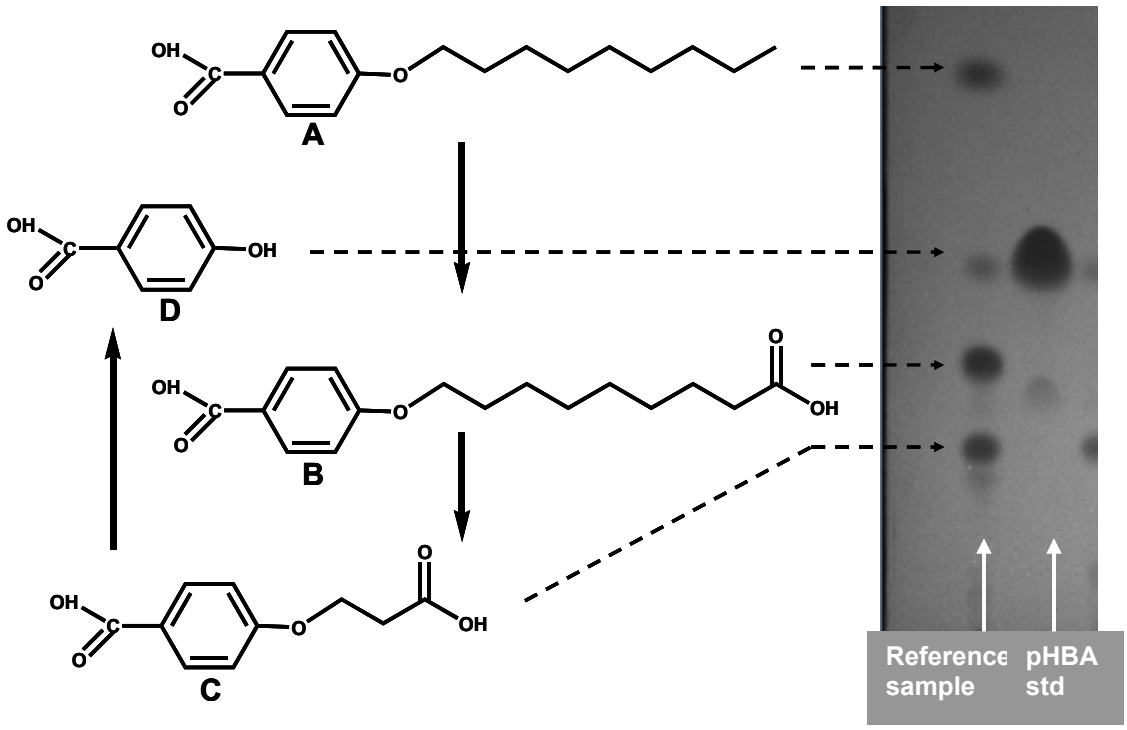
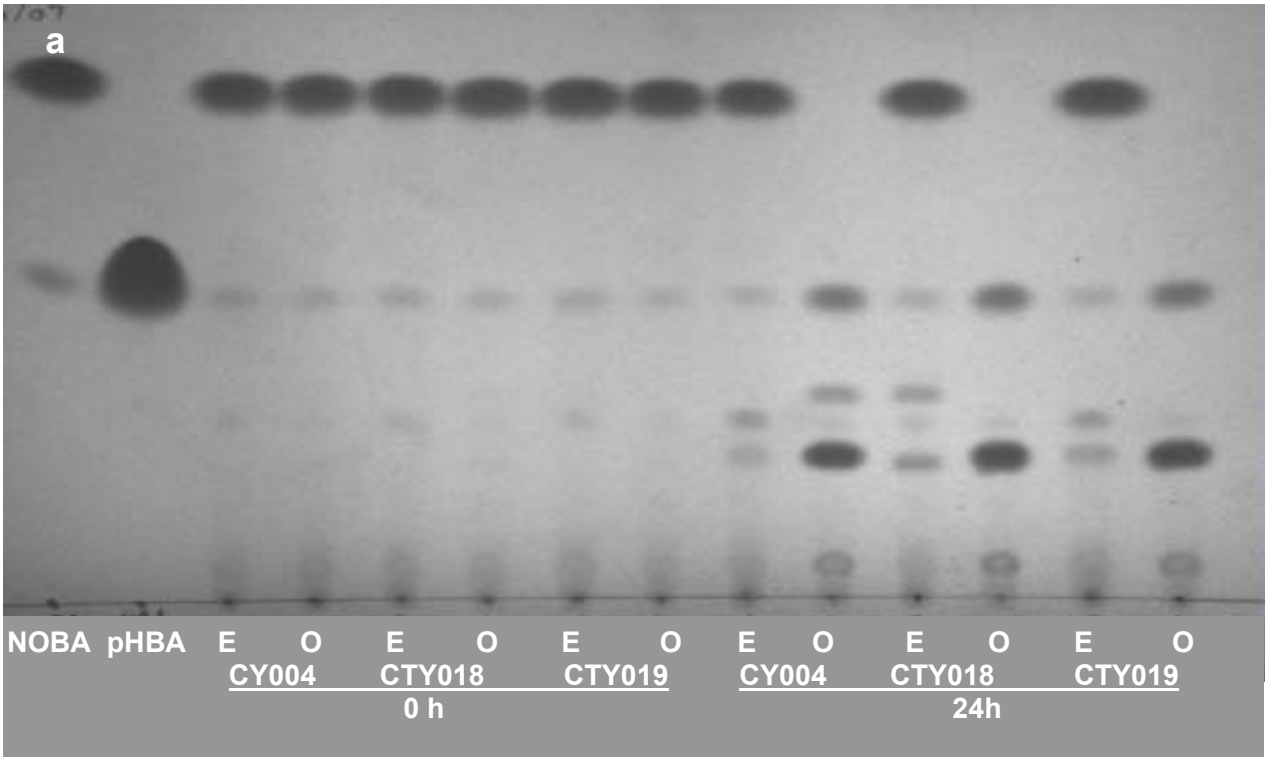


Figure 3.26: Conversion of NOBA to pHBA via its major intermediates identified by GC-MS analysis of intermediates isolated from TLC plates (Obiero, 2006). A – nonyloxybenzoic acid (NOBA), B – 4-(8-carboxyloctyloxy)benzoic acid, C – 4-(2-carboxyethoxy)benzoic acid, D – *p*-hydroxybenzoic acid (pHBA).

The NOBA, pHBA and intermediates are all readily detectable on appropriate TLC plates, allowing for rapid activity analysis. Due to the similarities between these two substrates, NOBA was chosen as a substrate for CYP102A1 biotransformations. Since the ether bonds of the pNCAs are cleaved by CYP102A1, a faster production of pHBA from NOBA was expected by strains expressing active CYP102A1. It was also possible that NOBA might be perceived as a fatty acid by CYP102A1 and be hydroxylated at subterminal positions.

TVN496 derivatives carrying single copies of the *CYP102A1* gene under the *POX2* and *ICL* promoters were compared for biotransformation of 4-nonyloxybenzoic acid (NOBA). Each strain was induced with ethanol and oleic acid (2% v/v) in separate flasks. The appropriate inducer (1% v/v) was supplemented after 24 hours, and then at 24 hour intervals thereafter (at 0.5% v/v). Six hours after the initial induction, NOBA (0.1% v/v) was added as substrate. Samples were collected regularly, and analysed using TLC (figure 3.27).



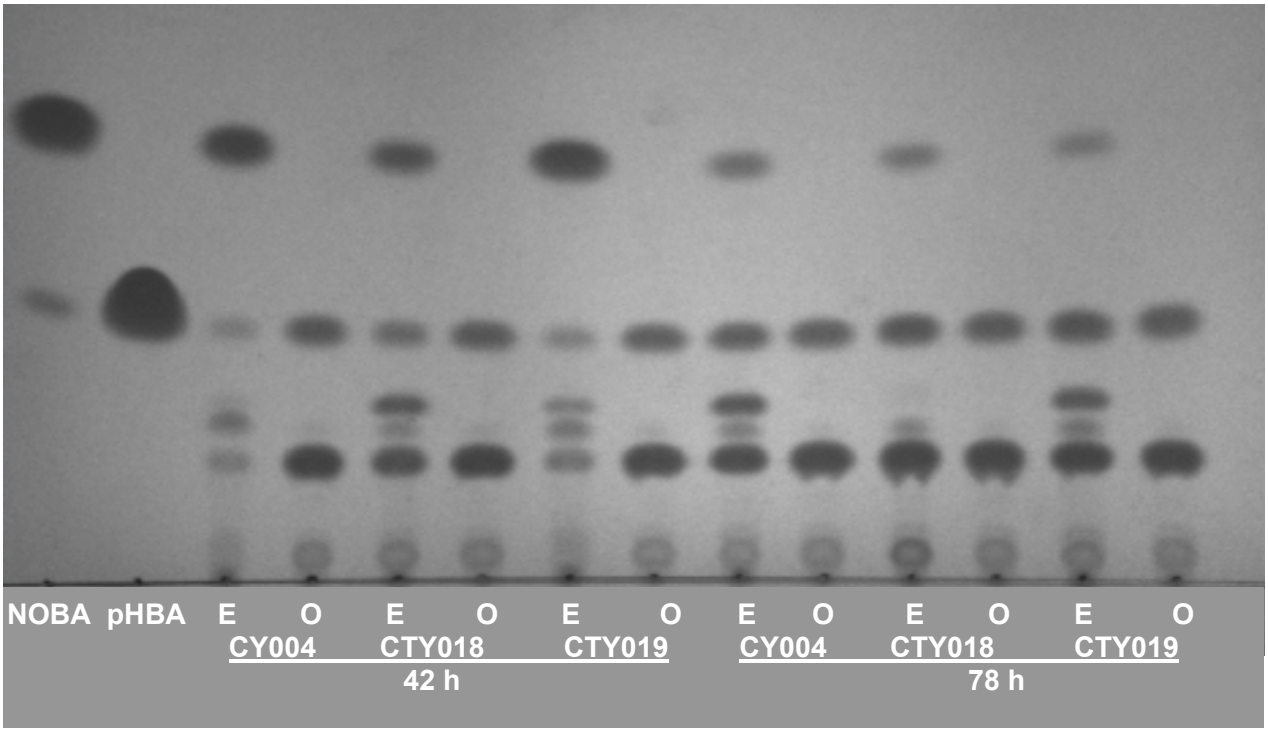


Figure 3.27: Biotransformation of 4-nonyloxybenzoic acid by TVN496 transformants containing *CYP102A1.* Strain CTY004 (negative control) CTY018 (*POX2*:CYP102A1) and CTY019 (*ICL:CYP102A1*) were induced with ethanol (E) or oleic acid (O) after 24 (2 % v/v), 48 (1 % v/v) and 72 (0.5 v/v) h growth. 4-Nonyloxybenzoic acid (NOBA) (0.1 % w/v) was added after 30 h. The final product was *para*-hydroxy benzoic acid (pHBA).

The small amount of pHBA present in the substrate standard is a result of spontaneous degradation of the substrate due to the instability of the nonyloxy-functional group, and accounts for the low amounts of pHBA observed in the 0 hour samples. It is clear from the TLC analysis that the strains induced with oleic acid rapidly within 24 h utilized all the substrate, regardless of whether or not the CYP gene was cloned. These results were similar to results previously obtained when *Y. lipolytica* strains containing multiple copies of *CYP53B1*, *CYP557A1* and *CYP52F1* under control of the *POX2* promoter were induced with oleic acid prior to addition of NOBA (Obiero, 2006). There was however after 24 h still a seemingly large amount of substrate present in the cultures induced with ethanol.

In fact, the substrate only started to significantly disappear after 42 hours, when the ethanol concentration assumingly was beginning to decline. Correspondingly, the pHBA production in these strains also only started catching up to the oleic acid induced strains after 42 hours.

The slight variations observed, (such as in the rates of degradation of the intermediates) were likely due to other differences between individual transformants, or small experimental variation. The bottom-line of these results is that the strains with cloned *CYP102A1* genes showed the same product patterns as the negative control strain. No unique bands (due to new sub-terminal hydroxylated products) were observed for them, nor did they appear to form more of the pHBA at an enhanced rate. This attempt to use whole cell biotransformations to distinguish between negative control strains and strains carrying the CYP102A1 gene was thus unsuccessful. The inability to detect CYP102A1 activity with this substrate might either be because NOBA and the corresponding dioic acid are not substrates for CYP102A1 or because the enzyme was not actively expressed. The trial did however reinforce the proposed inhibition of hydroxylase activity by ethanol.

FT-120 derivatives

FT-120 strains cloned with CYP102A1 were eligible for testing of fatty acid biotransformation, since fatty acids are the natural substrates of the enzyme, and these strains have their β-oxidation pathways completely disrupted. Subterminally hydroxylated products were expected as confirmation of CYP102A1 activity, since wild-type *Y. lipolytica* does not hydroxylate fatty acids at those positions.

After PCR confirmation of genomic DNA integration, strain CTY030, a derivative of FT-120 containing CYP102A1 under the control of the *ICL* promoter, was compared to strain CTY029, an FT-120 derivative without a cloned *CYP* gene, as

a negative control. Ethanol (2% v/v) was added as an inducer to 24 hour old main cultures, and was supplemented 24 hours after that (at 1% v/v), and at 24 hour intervals after that (at 0.5% v/v). The fatty acid substrates myristic acid and palmitic acid (0.5% w/v) were added 6 hours after the initial induction, and samples were collected at regular intervals.

GC analysis did not show formation of any unique products formed by the strains containing CYP102A1 (data not shown). While the NOBA used as a substrate for TVN496 derivatives may not be recognised as a substrate by CYP102A1, there is no question about its activity towards fatty acids in the range of C12-C22 (Noble *et al*, 1999). These results, together with those obtained for the TVN496 derivatives; indicate that there could be other factors influencing the activity of these enzymes.

A possible complication might be the soluble nature of the enzyme. Eukaryotic P450s generally tend to be membrane-bound, hence the hydrophobic substrates would accumulate in specific subcellular compartments, such as the ER and peroxisomes, where they would encounter all the enzymes involved in their metabolism. The CYP102A1 on the other hand, would most likely be localized in the cytosol, where it might never interact with the hydrophobic substrates. In such a case, a fully functional enzyme may be expressed by *Y. lipolytica* as a host, but it may be 'lost' in the cell in relation to the substrates. If this is the case whole cell assays of this yeast would never allow activity determination of these enzymes.

To confirm this notion, assays needed to be performed on cell-free extracts of the cells. Ideally, one should observe little difference in activity of microsomal fractions of the different strains (where the native P450s and their reductases will be located) but significantly increased activity in soluble fractions of strains in which CYP102A1 is present.

CPR assays based on the reduction of cytochrome *c* by the reductase component of the self-sufficient enzyme could be used on cell-free extracts of the CYP102A1 containing strains (Narhi and Fulco, 1987). In our research group, microsomal and soluble fractions were obtained from induced cells of the TVN496 and FT-120 derivative strains, based on the $CaCl_2$ precipitation method (Käppeli *et al*, 1982). Preliminary results obtained with several strains were very promising, as they convincingly demonstrated higher CPR activities in the soluble fractions of CYP102A1 containing strains compared with control strains. In one experiment the CPR activity in the soluble fraction of CTY029 (negative control) was 23 nmol.min^{-1}.mg protein^{-1}, while the CPR activity in the soluble fraction of the test strain CTY030 (*ICL:CYP102A1*) was 775 nmol.min^{-1}.mg protein^{-1} after induction with ethanol (Mabwe, personal communication). It would be highly unlikely that the reductase activity would not be accompanied by complementary hydroxylase activity, indicating that the CYP102A1 enzyme is indeed successfully expressed. These results fundamentally confirm the concept that there exists a lack of the necessary enzyme-substrate interaction within the *Y. lipolytica* whole cells when the hydroxylase is a soluble enzyme.

Chapter 4: Concluding remarks

The cytochrome P450 monooxygenases are a diverse group of enzymes capable of hydroxylating a range of hydrophobic substrates. These enzymes hold several advantages over chemical syntheses, especially superior regio-, enantio-, and stereoselectivities; and far milder operating conditions with regard to pH, temperature and pressure. They therefore have excellent potential as alternatives to traditional chemistry procedures for industrial applications involving conversion of inexpensive substrates into higher value products. Large scale applications of these enzymes have however been hindered by their mandatory requirement for regeneration of expensive cofactors, such as NAD(P)H.

The best solution to this problem seems to be the use of whole cells, where the enzymes can function in their natural environment. Often the natural hosts are incapable of managing the hydrophobic substrates for the P450 reactions or it is difficult to optimize conditions for P450 production and biotransformation. Heterologous expression in well characterized hosts is an attractive alternative Many P450s have been successfully expressed in a range of bacterial and yeast hosts. However, the enzymes have mostly been characterized in cell free extracts or as isolated enzymes. This generally requires co-expression of a suitable cytochrome P450 reductase (either as a separate of fused enzymes) or reconstitution with a suitable reductase produced separately.

Yarrowia lipolytica is an attractive eukaryotic host for expression of P450 enzymes, and has been successfully applied for this purpose. A large number of genetic tools are available for optimization of heterologous expression, including strain-specific means of integration, strong promoters to control gene expression, and marker genes allowing multiple integrations of vector DNA into the host genome (Juretzek *et al*, 2001; Madzak *et al*, 2004).

In our research group, two heterologous fungal P450 genes, *CYP53B1* and *CYP557A1*, have been expressed in *Y. lipolytica* under the control of the *POX2* promoter which is induced by fatty acids. In the case of CYP53B1 benzoate-*para*-hydroxylase activity was detectable, but disappointingly low, probably because the fatty acids used for induction were also inducers and subsrates of the wild-type alkane and fatty acid hydroxylases of *Y. lipolytica*, encoded by twelve *CYP52* genes (*CYP52F1 – CYP52F12)* (Fickers *et al*, 2005). Problems were thus also encountered with the putative alkane and fatty acid hydroxylase encoded by *CYP557A1.* In this case the anticipated activity of the expected gene product resembles the activities of the wild-type alkane and fatty acid hydroxylases. Since the inducers of the *POX2* promoter also induced the wild-type P450s, differentiation between the activities of the cloned and native P450s became difficult. The determination of cloned activities was further limited by the lack of suitable negative control strains. In the current study, we thus set out to clone P450 encoding genes under the control of the *ICL* promoter instead. This promoter is inducible by ethanol, which in previous studies had delayed the induction of wild-type P450 activity (Obiero, 2005; Van Rooyen, 2006, Shiningavamwe *et al*, 2006).

Vectors allowing the expression of P450s under the regulation of the *ICL* promoter were successfully constructed. These vectors are equipped with zeta elements for either homologous or heterologous integration into host genomes, depending on the host strain used. Versions of these vectors containing the defective ura3d4 marker were also constructed to allow multiple copies of the cassette to be integrated into the host genome.

These vectors were cloned into two *Y lipolytica* strains, namely TVN496 and FT120. TVN496 had previously been derived from E150 by cloning of an additional *YlCPR* under control of the *ICL* promoter (Setati, unpublished results). TVN496 contains zeta elements since it had been derived from E150, a zeta containing strain from the French inbreeding line. FT120 is a β-oxidation blocked

strain derived from another parental strain Po1d, which does not contain zeta elements. These vectors allowed successful integration into the host genomes, with transformation efficiencies comparable to their *POX2* containing counterparts, for strains with and without zeta elements. More relevant negative control strains were also constructed for biotransformation studies. Strains TVN496, FT120 and E150 were transformed with versions of the expression vectors lacking *YlCPR* and *CYP* genes, for restoration of the prototrophic phenotypes. Introduction of cloned vectors were confirmed with PCR and a number of biotransformation experiments were carried out to confirm expression and activity of the cloned genes. These biotransformation experiments were carried out in test tubes or small shake flasks and conditions were not optimized. These results are therefore preliminary and should be treated with caution. A few common trends did however emerge.

Enzyme activity of the benzoate-*para*-hydroxylase encoding *CYP53*, under the regulation of the *ICL* promoter, was shown to compare favorably with, and in some cases exceed, the activity driven by the *POX2* promoter. In a whole cell biotransformation experiment with a strain expressing *CYP53B1* under control of the *ICL* promoter, induction with oleic acid gave better results than induction with ethanol. After oleic acid induction this strain also exhibited higher activity than the best strain from previous studies expressing a single copy of *CYP53B1* under control of the *POX2* promoter, and comparable activity to the TVN91 strain which contains multiple copies of *POX2:CYP53*. Since cell-free extracts from *Y. lipolytica* showed comparable heterologous β-galactosidase activities under the ICL promoter with either oleic acid or ethanol induction (Juretzek *et al*, 2000a), we concluded that the ethanol has an inhibitory effect at another point in the process of hydrophobic substrate metabolism by whole cells. These results correspond to results obtained for heterologous expression of bovine P45017α in *Y. lipolytica* under the ICL promoter. This study by Juretzek *et al*. (2000b) showed that microsomal hydroxylase activities differed only marginally, whether ethanol or hexadecane was used as an inducer. In sharp contrast to this, the

whole cell activity, after hexadecane induction, was significantly higher than activity induced by ethanol. We concluded that the ethanol is likely repressing enzymes involved in the uptake and/ or sub-cellular transport of hydrophobic substrates within the cells, which would definitely be switched on when oleic acid is used as an inducer.

In the case of strains with the CYP557A1 gene cloned under control of the ICL promoter, we have in a few cases (during this study and in follow up experiments in our group) obtained results which indicated that CYP557A1 was functionally expressed. However, these results are not conclusive and we require experiments with microsomal fractions to confirm increased P450 content with carbon monoxide difference spectra and to do activity assays. Results obtained with the *CYP557* containing strains support the suggestion that ethanol is not allowing sufficient uptake and/ or transport of the substrates within the cells, since the activities obtained were generally higher when ethanol was added only once, than when ethanol was added at regular intervals.

The activities for CYP53 and CYP557 reached higher levels when the ethanol levels subsided, and in cases where ethanol was only added once rather than continually. Therefore it is clear that the strategy of continuous addition of ethanol for simultaneous induction of cloned P450s and repression of native P450s, leads to a continual repression of the general ability of the cell to deal with these kinds of substrates, resulting in an overall reduced activity.

Preliminary CPR assays carried out in our group with separated soluble and microsomal fractions have indicated that the CYP102A1 gene was functionally expressed by *Y. lipolytica* under both the ICL and POX2 promoters (Mabwe, personal communication). Increased CPR activity was observed in the soluble fractions of strains containing the *CYP102A1* gene after induction with the appropriate inducers. The activity of the cloned CYP102A1 could not be detected using whole cells however, even when oleic acid was used as an

inducer. This was presumably due to the absence of interaction between the soluble enzymes which were confined to the cytosol, and the hydrophobic substrates, which were localized in specific organelles (ER and/or peroxisomes) within the cells.

It is acknowledged that the characterization of the strains engineered in this study are thus far still very preliminary, and more thorough work in this regard is still outstanding. We now have a range of strains which thus far appear to be functionally expressing the three discussed P450s, and relevant control strains to compare them to. The results for the CYP53 showed us that many transformants need to be screened to identify the most promising candidate. While only the best performers from preliminary screens were assigned CTY names so far, more transformants are also still available for further evaluation. There is however an urgent requirement for more appropriate screening methods for some of the enzymes. Further investigations need to be carried out to identify the optimal conditions required for the activities of these enzymes, before the other transformants can be properly evaluated.

Copy numbers for the strains still need to be confirmed before more accurate conclusions can be drawn about their relative activities. More extensive assays on the appropriate cell free extracts can be performed for the different enzymes, to test for any differences between the ICL and POX2 promoters. The use of ethanol for the induction of the ICL promoter has proven to be detrimental to activity when whole cells are used. The hindrance is aggravated when ethanol is continually fed during the biotransformation. This means that unless another substance for use as a selective inducer of the ICL promoter can be identified, the ICL promoter can not be successfully applied for whole cell biocatalysis where background activities need to be avoided. The use of other inducer-substrate combinations, such as acetate as inducer and butylbenzene as substrate are currently under investigation in our research group.

Y. lipolytica can possibly be used as a host for the production of active *CYP102A1*, since soluble fractions displayed CPR activity that can be ascribed to CYP102A1. The application of whole cells expressing the soluble CYP102A1 for biotransformations, is not viable however, because hydrophobic substrates are apparently chanelled to the endoplasmic reticulum where the endogenous P450s are located.

The membrane-bound eukaryotic counterpart of CYP102A1 from the fungus *Fusarium oxysporum*, CYP505 (P450foxy), is an attractive candidate for heterologous expression by *Y. lipolytica*. Expression of these P450s in *E. coli* and *S. cerevisiae* resulted in soluble versions of the protein, due to inadequate post-translational modifications performed by these cells (Kitazume *et al*, 2000; Kitazume *et al*, 2002). *Y. lipolytica* has already shown very accurate post-translational modification mimicry of natural hosts, when the heterologous expression of a range of different proteins by this yeast was tested (Muller *et al*, 1998). Therefore, we don't forecast these kinds of problems if *Y. lipolytica* is to be used as a host for the expression of self-sufficient P450s from fungal origin. If the expressed enzymes are localized in the membranes, they will reap all the benefits of *Y. lipolytica* as a host for P450s in terms of substrate uptake and transport, and the efficient regeneration of the NADPH required by these enzymes. It will be very interesting to see whether these very efficient subterminal fatty acid hydroxylases would out-compete the endogenous P450s in *Y. lipolytica,* producing hydroxy fatty acids rather than dioic acids in strains with disrupted β-oxidation.

However, ultimately we aim to develop a system that will avoid expression of the endogenous P450s, while still making maximum use of the specialized hydrophobic substrate uptake system and the co-factor regeneration system that is already present in *Y. lipolytica*. This might eventually require deletion of some of the highly expressed *CYP52s* in *Y. lipolytica*. Such a system will allow the use of whole cells of *Y. lipolytica* strains expressing a range of *CYP* genes from

animal, plant or fungal origin for the hydroxylation of different hydrophobic substrates.

List of references

Ayala, M., Torres, E. (2004). Enzymatic activation of alkanes: constraints and perspective. *Applied catalysis A: General*. **272:** 1-13.

Backes, W., Kelley, R.W. (2003). Organization of multiple cytochrome P450s with NADPH-cytochrome P450 reductase in membranes. *Pharmacology and therapeutics.* **98:** 221-233.

Barth, G., Gaillardin, C. (1996). *Yarrowia lipolytica.* In *Non conventional yeasts in biotechnology*. Wolf K. (ed). Springer-verlang; 313-388.

Bourel, G., Nicaud, J-M., Nthangeni, B., Santiago-Gomez, P., Belin, J-M., Husson, F. (2004). Fatty acid hydroperoxide lyase from green bell pepper: cloning in *Yarrowia lipolytica* and biogenesis of volatile aldehydes. *Enzyme and microbial technology.* **35:** 293-299.

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

Butte, W. (1983). Rapid method for the determination of fatty acid profiles from fats and oils using trimethylsulfonium hydroxide for transesterification. *Journal of chromatography.* **261:** 142-145.

Casaregola, S., Feynerol, C., Diez, M., Fournier, P., Gaillardin, C. (1997). Genomic oranization of the yeast *Yarrowia lipolytica. Chromasoma.* **106:** 380-390.

Cereghino, J.L., Cregg, J.M. (2000). Heterologous expression in the methylotrophic yeast *Pichia pastoris. FEMS microbiology reviews.* **24:** 45-66.

Cereghino, G.P.L., Cereghino, J.L., Ilgen, C., Cregg, J.M. (2002). Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris. Current opinion in biotechnology.* **13:** 329-332.

Cheng, J., Wan, D-F., Gu, J.R., Gong, Y., Yang, S-L., Hao, D-C., Yang, L. (2006). Establishment of a yeast system that stably expresses human cytochrome P450 reductase: application for the study of drug metabolism of cytochrome P450s *in vitro. Protein expression and purification.* **47:** 467-476.

Cordero Otero, R., Gaillardin, C. (1996). Efficient selection of hygromycin-B-resistant *Yarrowia lipolytica* transformants. *Applied microbiology and biotechnology*. **46(2):** 143-148.

Costas, M., Chen, K., Que, L. Jr. (2000). Biomimetic nonheme iron catalysts for alkane hydroxylation. *Coordination chemistry reviews.* **200-202:** 517-544.

De Mot, R., Parret, A.H.A. (2002). A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes. *Trends in microbiology.* **10(11):** 502-508.

Duetz, W.A., Van Beilen, J.B., Witholt, B. (2001). Using proteins in their natural environment: potential and limitations of microbial whole-cell hydroxylations in applied biocatalysis. *Current opinion in biotechnology.* **12:** 419-425.

Farinas, E.T., Alcalde, M., Arnold, F. (2004). Alkene epoxidation catalyzed by cytochrome P450 BM-3 139-3. *Tetrahedron.* **60(3):** 525-528.

Fickers, P., Le Dall, M.T., Gaillardin, C., Thonart, P., Nicaud, J.M. (2003). New disruption cassettes for rapid gene disruption and marker rescue in the yeast *Yarrowia lipolytica*. *Journal of microbiological methods*. **55:** 727-737.

Fickers, P., Benetti, P-H., Wache, Y., Marty, A., Mauersberger, S., Smit, M.S., Nicaud, J-M. (2005). Hydrophobic substrate utilisation by the yeast *Yarrowia lipolytica*, and its potential applications. *FEMS yeast research.* **5:** 527-543.

Glieder, A., Farinas,E.T., Arnold, F.H., (2002). Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nature biotechnology*. **20:** 1135-1139.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of molecular biology.* **166(4):** 557-580.

Hanley, K., Nguyen, L.V., Khan, F., Pogue, G.P., Vojdani, F., Panda, S., Pinot, F., Oriedo, V.B., Rasochova, L., Subramanian, M., Miller, B., White, E.L. (2003). Development of a plant viral-vector-based gene expression assay for the screening of yeast cytochrome P450 monooxygenases. *Assay and drug development technologies.***1(1-2):** 147-160.

Hannemann, F., Bichet, A., Ewen, K.M., Bernhardt, R. (2007). Cytochrome P450 systems – biological variations of electron transport chains. *Biochimica et biophysica acta.* **1770(3):** 330-344.

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

Juretzek, T., Wang, H-J., Nicaud, J-M., Mauersberger, S., Barth, G. (2000). Comparison of promoters suitable for regulated overexpression of β-galactosidase in the alkane assimilating yeast *Yarrowia lipolytica. Biotechnological bioprocess engineering.* **5:** 320-326.

Juretzek, T., Mauersberger, S., Barth, G. (2000). Rekombinante haploide oder diploide *Yarrowia lipolytica* Zellen zur funktionellen heterologen Expression von Cytochrom P450 Systemen. *German patent*, DE 198 30 905.8, DE 199 32 811.0, PCT/DE 99/02174.

Juretzek, T., Le Dall, M-T., Mauersberger, S., Gaillardin, C., Barth, G., Nicaud, J-M. (2001). Vectors for gene expression and amplification in the yeast *Yarrowia lipolytica*. *Yeast.* **18:** 97-113.

Käppeli, O., Sauer, M., Fiechter, A. (1982) Convenient procedure for the isolation of highly enriched, cytochrome p-450-containing microsomal fraction from *Candida tropicalis*. *Analytical Biochemistry.* **126(1):** 179-182.

Kelly, S.L., Lamb, D.C., Jackson, C.J., Warrilow, A.G.S., Kelly, D.E. (2003). The biodiversity of microbial cytochromes P450. *Advances in microbial physiology.* **47:** 131-186.

Kitazume, T., Takaya, N., Nakayama, N., Shoun, H. (2000). *Fusarium oxysporum* fatty-acid subterminal hydroxylase (CYP505) is a membrane-bound eukaryotic counterpart of *Bacillus megaterium* cytochrome P450BM3. *The journal of biological chemistry*. **275(50):** 39734-39740.

Kitazume, T., Tanaka, A., Matsuyama, S., Takaya, N., Shoun, H. (2002). Analysis of fungal P450foxy (*CYP505*), a fused protein of cytochrome P450 and its reductase. *International congress series*. **1233:** 143-149.

Kitazume, T., Tanaka, A., Takaya, N., Nakamura, A., Matsuyama, S., Suzuki, T., Shoun, H. (2002). Kinetic analysis of hydroxylation of saturated fatty acids by recombinant P450foxy produced by an *Escherichia coli* expression system. *European journal of biochemistry / FEBS.* **269:** 2075-2082.

Kolar, N.W., Swart, A.C., Mason, J.I., Swart, P. (2007). Functional expression and characterization of human cytochrome P45017α in *Pichia pastoris. Journal of biotechnology.* **129(4):** 635-644.

Le Dall, M-T., Nicaud, J.M., Gaillardin, C. (1994). Multi-copy integration in the yeast *Yarrowia lipolytica. Current genetics.* **26:** 38-44.

Lentz, O., Urlacher, V., Schmid, R.D. (2004). Substrate specificity of native and mutated cytochrome P450 (CYP102A3) from *Bacillus subtilis. Journal of biotechnology.* **108:** 41-49.

Lida, T., Sumita, T., Ohta, A., Takagi, M. (2000). The cytochrome P450ALK mmultigene family of an *n*-alkane assimilating yeast *Yarrowia lipolytica*: cloning and characterization of genes coding for new CYP52 family members. *Yeast.* **16:** 1077-1087.

Madzak, C., Treton, B., Blanchin-Roland, S. (2000). Strong hybrid promoters and integrative expression/ secretion vectors for quasi-constitutive expression of heterologous proteins in the yeasy *Yarrowia lipolytica. Journal of molecular microbiology and biotechnology.* **2(2):** 207-216.

Madzak, C., Gaillardin, C., Beckerich, J-M. (2004). Heterologous protein expression and secretion in the non-conventional yeast *Yarrowia lipolytica*: a review. *Journal of biotechnology*. **109:** 63-81.

Mauersberger, S., Drechsler, H., Oehme, G., Muller, H-G. (1992). Substrate specificity and stereoselectivity of fatty alcohol oxidase from the yeast *Candida maltosa. Applied microbiology and biotechnology.* **37:** 66-73.

Mauersberger, S., Wang, H-J., Gaillardin, C., Barth, G., Nicaud, J-M. (2001). Insertional mutagenesis in the *n*-alkane assimilating yeast *Yarrowia lipolytica*:

generation of tagged mutations in genes involved in hydrophobic substrate utilization. *Journal of bacteriology*. **183(17):** 5102-5109.

Maurer, S.C., Kuhnel,, K., Kaysser, L.A., Eiben, S., Schmid, R., Urlacher, V. (2005). Catalytic hydroxylation in biphasic systems using CYP102A1 mutants. *Advanced synthesis and catalysis*. **347:** 1090-1098.

Miura, Y., Fulco, A.J. (1974). (ω - 2) Hydroxylation of fatty acids by a soluble system from *Bacillus megaterium. The journal of biological chemistry.* **249(6):** 1880-1888.

Mlickova, K., Luo, Y., d'Andrea, S., Pec, P., Chardot, T., Nicaud, J-M. (2004). Acyl-CoA oxidase, a key step in for lipid accumulation in the yeast *Yarrowia lipolytica. Journal of molecular catalysis B: Enzymatic.* **28:** 81-85.

Moe, L.A., Hu, Z., Deng, D., Austin, R.N., Groves, J.T., Fox, B.G. (2004). Remarkable aliphatic hydroxylation by the diiron enzyme toluene 4-monoxygenase in reactions with radical or cation diagnostic probes norcarane, 1,1-dimethylcyclopropane, and 1,1-diethylcycloprpane. *Biochemistry.* **43:** 15688-15701.

Muller, S., Sandal,T., Kamp-Hansen, P., Dalbøge, H. (1998). Comparison of expressionsystems in the yeasts *Saccharomyces cerevisiae, Hansenula polymorpha, Klyveromyces lactis, Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Cloning of two novel promoters from *Yarrowia lipolytica*. *Yeast.* **14:** 1267-1283.

Murakami, H., Yabusaki, Y., Sakaki, T., Shibata, M., Ohkawa, H. (1990). Expression of cloned yeast NADPH-cytochrome P450 reductase gene in *Saccharomyces cerevisiae. Journal of biochemistry.* **108:** 859-865.

Nakayama, N., Takemae, A., Shoun, H. (1996). Cytochrome P450foxy, a catalytically self-sufficient fatty acid hydroxylase of the fungus *Fusarium oxysporum. Journal of biochemistry.* **119:** 435-440.

Narhi, L.O., Fulco, A.J. (1986). Characterization of a catalytically self-sufficient 119-000-Dalton cytochrome P450 monooxygenase induced by barbituates in *Bacillus megaterium. The journal of biological chemistry.* **261(16):** 7160-7169.

Narhi, L.O., Fulco, A.J. (1987). Identification and characterization of two functional domains in cytochrome P-450BM-3, a catalytically self-sufficient monooxygenase induced by barbituates in *Bacillus megaterium. The journal of biological chemistry.* **262(14):** 6683-6690.

Narhi, L.O., Wen, L-P., Fulco, A.J. (1988). Characterization of the protein expressed in *Escherichia coli* by a recombinant plasmid containing the *Bacillus megatereium* cytochrome P-450BM-3 gene. *Molecular and cellular biology.* **79:** 63-71.

Nicaud, J-M., Madzak, C., Van den Broek, P., Gysler, C., Duboc, P., Niederberger, P., Gaillardin, C. (2002). Protein expression and secretion in the yeast *Yarrowia lipolytaica*. *FEMS yeast research.* **2(3):** 371-379.

Noble, M.A., Miles,C.S., Chapman, S.K., Lysek, D.A., Mackay, A.C., Reid, G.A., Hanzlik, R.P., Munro, A.W. (1999). Roles of key active site residues in flavocytochrome P450 BM3. *The biochemical journal.* **339:** 371-379.

Nthangeni, M.B., Urban, P., Pompon, D., Smit, M.S., Nicaud, J-M. (2004). The use of *Yarrowia lipolytica* for the expression of human cytochrome P450 CYP1A1. *Yeast.* **21(7):** 583-592.

Obiero, G.O. (2006). Evaluation of *Yarrowia lipolytica* as a host for cytochrome P450 monooxygenase expression. (PhD thesis). University of the Free State.

Peters, M.W., Meinhold, P., Glieder, A., Arnold, F.H. (2003). Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3. *Journal of the American chemical society.* **125(44):** 13442-13450.

Picataggio, S., Rohrer, T., Deanda, K., Lanning, D., Reynolds, R., Mielenz, J., Eirich, L.D. (1992). Metabolic engineering of *Candida tropicalis* for the production of long-chain fatty dicarboxylic acids. *Biotechnology.* **10:** 894-898.

Pignede, G., Wang, H-J., Fudalej, F., Seman, M., Gaillardin, C., Nicaud, J-M. (2000). Autocloning and amplification of *LIP2* in *Yarrowia lipolytica*. *Applied and environmental microbiology.* **66 (8):** 3283-3289.

Roberts, G.A., Grogan, G., Greter, A., Flitsch, S.L., Turner, N.J. (2002). Identification of a new class of cytochrome P450 from a *Rhodococcus* sp. *Journal of bacteriology.* **184(14):** 3898-3908.

Ruettinger, R.T., Fulco, A.J. (1981). Epoxidation of unsaturated fatty acids by a soluble cytochrome P-450-dependant system from *Bacillus megaterium. The journal of biological chemistry.* **256(11):** 5728-5734.

Sakaki,T., Inouye, K. (2000). Practical application of mammalian cytochrome P450. *Journal of bioscience and bioengineering.* **90(6):** 583-590.

Sambrook, J., Fritsch, E.F., Maniates, F., (1989). Molecular cloning: a labotatory manual, 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Scheller, U., Zimmer, T., Becher, D., Schauer, F., Schunck, W-H. (1998). Oxygenation cascade in the conversion of *n*-alkanes to α,*ω*-dioic acids catalyzed by cytochrome P450 52A3. *The journal of biological chemistry.* **273(49):** 32528-32534.

Schneider, S., Wubbolts, M.G., Oesterhelt, G., Sanglard, D., Witholt, B. (1999). Controlled regioselectivity of fatty acid oxidation by whole cells producing cytochrome P450BM-3 monooxygenase under varied dissolved oxygen concentrations. *Biotechnnology and bioengineering*. **64(3):** 333-341.

Shoun, H., Sudo, Y., Beppu, T. (1985). Subterminal hydroxylation of fatty acods by a cytochrome P-450-dependant enzyme system from a fungus, *Fusarium oxysporum. Journal of biochemistry.* **97:** 755-763.

Schwaneberg, U., Schmidt-Dannert, C., Schmitt, J., Schmid, R.D. (1999). A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A. *Analytical biochemistry.* **269:** 359-366.

Schwaneberg, U., Appel, D., Schmitt, J., Schmid, R.D. (2000). P450 in biotechnology: zinc driven ω-hydroxylation of *p*-nitrophenoxydodecanoic acid using P450 BM-3 F87A as a catalyst. *Journal of biotechnology*. **84:** 249-257.

Shiningavamwe, A.N. (2004). Molecular cloning and expression of cytochrome P-450 monooxygenases from *Rhodoturula* spp. in *Yarrowia lipolytica*. (PhD thesis). University of the Free State.

Shiningavamwe, A., Obiero, G., Albertyn, J., Nicaud, J-M., Smit, M. (2006). Heterologous expression of the benzoate-*para*-hydroxylase encoding gene (*CYP53B1*) from *Rhodoturula minuta* by *Yarrowia lipolytica*. *Applied microbiology and biotechnology*. **72:** 323-329.

Smit, M.S., Mokgoro, M.M., Setati, E., Nicaud, J-M. (2005). α,ω-dicarboxylic acid accumulation by acyl-CoA oxidase deficient mutants of *Yarrowia lipolytica*. *Biotechnology letters.* **27:** 859-864.

Swennen, D., Paul, M-F., Vernis, L., Beckerich, J-M., Fournier, A., Gaillardin, C. (2002). Secretion of active anti-Ras single-chain Fv antibody by the yeasts *Yarrowia lipolytica* and *Kluyveromyces lactis. Microbiology.* **148:** 41-50.

Thevenieau, F., Le Dall, M-T., Nthangeni, B., Mauersberger, S., Marchall, R., Nicaud, J-M. (2007). Characterization of *Yarrowia lipolytica* mutants affected in hydrophobic substrate utilization. *Fungal genetics and biology.* **6:** 531-542.

Truan, C., Cullin, C., Reisdorf, P., Urban, P., Pompon, D. (1993). Enhanced *in vivo* activities of mammalian P450s in engineered yeast cells producing high levels of NADPH-P450 reductase and human cytochrome b5. *Gene.* **125(1):** 49-55.

Urlacher, V.B., Lutz-Wahl, S. Schmid, R.D. (2004). Microbial P450 enzymes in biotechnology. *Applied microbiology and biotechnology*. **64:** 317-325.

Van Beilen, J.B., Holtackers, R., Luscher, D., Bauer, U., Witholt, B., Duetz, W.A. (2005). Biocatalytic production of perillyl alcohol from limonene by using a novel *Mycobacterium* sp. cytochrome P450 alkane hydroxylase expressed in *Pseudomonas putida*. *Applied and environmental microbiology.* **71(4):** 1737-1744.

Van Beilen, J.B., Funhoff, E.G. (2005). Expanding the alkane oxygenase toolbox: new enzymes and applications. *Current opinion in biotechnology.* **16:** 308-314.

Van Beilen, J.B., Funhoff, E.G. (2007). Alkane hydroxylases involved in microbial alkane degradation. *Applied microbiology and biotechnology*. **74:** 13-21.

Van Dyk, M.S., Van Rensburg, E., Moleleki, N. (1998). Hydroxylation of (+) limonene, (-) α-pinene and (-) β-pinene by a *Hormonema* sp. *biotechnology letters.* **20 (4):** 431-436.

Van Rooyen, N. (2005). Biotransformation of alkanes, alkylbenzenes and their derivatives by genetically engineered *Yarrowia lipolytica*. (M.Sc dissertation). University of the Free State.

Werck-Reichart, D., Feyereisen, R. (2000). Cytochrome P450: a success story. *Genome biology.* **1(6):** reviews3003.1-3003.9.

Wong, L-L. (1998). Cytochrome P450 monooxygenases. *Current opinion in chemical biology.* **2:** 263-268.

http://drnelson.utmem.edu/CytochromeP450.html

http://www.chem.ed.ac.uk/chapman/p450.html

http://cbi.labri.fr/Genolevures/elt/YALI

http://www.ncbi.nlm.nih.gov/BLAST/

http://www.ncbi.nlm.nih.gov/

www.fermentas.com/doubledigest/index.html

http://www.ebi.ac.uk/clustalw/index.html

http://www.yeastern.com

Abstract

Cytochrome P450 monooxygenases are enzymes capable of efficiently hydroxylating hydrophobic substrates with high regio-, stereo-, and enantioselectivies; under mild reaction conditions. They are therefore attractive alternatives to traditional chemistry for the synthesis of hydroxylated products. Their use in large scale applications has however been hindered by their requirement for continuous supply of reducing cofactors. Self-sufficient P450s occur, containing fused hydroxylase and reductase domains in a single polypeptide. They exhibit higher activities than any other reported P450s.

Yarrowia lipolytica is a yeast capable of efficient degradation of hydrophobic substrates and growth on alkanes as sole carbon source. The intricate pathways involved in hydrophobic substrate metabolism within this yeast involve initial hydroxylation by P450s. *Y. lipolytica* contains 12 alkane and fatty acid hydroxlyase encoding genes *CYP52F1* to *CYP52F12*. This yeast has been widely tested for biotransformations of inexpensive substrates to more valuable products. *Y. lipolytica* has also been investigated as a host for the heterologous expression of diverse foreign proteins, including cytochrome P450s. Various genetic tools are available for this purpose, including strong, inducible promoters, (e.g. POX2 and ICL); the defective selection marker ura3d4 for integration of cloned genes in higher copy numbers; specific targets for integration; and customized host strains.

Heterologous expression of two *CYP* genes, by *Y.lipolytica* has previously been investigated in our research group. These were *CYP53B1*, encoding a benzoate-*para*-hydroxylase, and *CYP557A1,* a putative alkane and fatty acid hydroxylase. They were cloned in single and multiple copies under the control of the *POX2* promoter. Problems were encountered with the *POX2* promoter when whole cells were used for biotransformations, since native P450s were also induced by the fatty acids used as inducers for the *POX2* promoter, and thus interfered with

activity determination of cloned P450s. A further limitation to this study was the lack of appropriate negative control strains for accurate comparisons.

In the current study, the same P450s were cloned into *Y. lipolytica* under the control of the *ICL* promoter. More appropriate negative controls strains were also constructed. Both ethanol and oleic acid can induce the *ICL* promoter. Ethanol induction of the *CYP53* expressing strains resulted in lower whole cell activities than oleic acid induction. It has been reported in the literature that activities of enzymes expressed under the *ICL* promoter in cell-free extracts from cells induced with either oleic acid or ethanol were comparable. It therefore seemed that ethanol was repressing other components of the metabolic pathways for the degradation of hydrophobic substrates in the cells.

Additionally no convincing activity was discernable for the *ICL* regulated CYP557 containing strains compared to the controls under *POX2* regulation, and negative control strains. Continuous ethanol addition again led to reduced activities compared to situations where ethanol was added only once, confirming the metabolic inhibition by ethanol.

It can be concluded that ethanol induction of the *ICL* promoter is not viable for whole cell biotransformations of hydrophobic substrates. Different inducers which will repress induction of the endogenous *CYP* genes but will allow induction of hydrophobic substrate uptake systems must be identified if the *ICL* promoter is to be used for the expression of cloned *CYP* genes.

The *CYP102A1* gene encoding the self-sufficient subterminal fatty acid hydroxylase from *Bacillus megaterium* was also cloned into *Y. lipolytica* under the *POX2* and *ICL* promoters. Strains containing this gene did not display detectable subterminal hydroxylation of fatty acids or 4-nonyloxy-benzoic acid which was used as a fatty acid analogue. Significantly increased cytochrome *c* reductase activities were however detected in the soluble fraction of cell free

extracts from test strains compared to the negative control strains. Microsomal activities were comparable in cells with and without cloned *CYP102A1*. This indicated that the *CYP102A1* was functionally expressed, but located in the soluble fraction, where it was unlikely to interact with the hydrophobic substrates. It was therefore concluded that *Y. lipolytica* is not a suitable host for whole cell biotransformations using the CYP102A1 enzyme.

Key terms: *Yarrowia lipolytica*, heterologous expression, cytochrome P450 monooxygenase, ICL promoter, substrate hydroxylation, hydrophobic substrates, whole cell biotransformations, benzoate-*para*-hydroxylase, self-sufficient P450s, CYP102A1.