FACULTY OF SCIENCES

A CYP solution: an enzyme-based process for the production of long-chain α,ω-dicarboxylic acids

Ir. Delphine Devriese

Supervisor:
Prof. Dr. Bart Devreese

Academic year 2020–2021

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Doctor of Sciences: Biochemistry and Biotechnology

GHENT UNIVERSITY

LM-UGent

Supervisor

Prof. Dr. Bart Devreese
Laboratory of Microbiology – Protein Research Unit, Ghent University

Examination committee

Prof. Dr. Savvas Savvides (Chair)
VIB-UGent Center for Inflammation Research

Dr. Kenneth Verstraete (Secretary)
VIB-UGent Center for Inflammation Research

Prof. Dr. ir. Inge Van Bogaert
Centre for Synthetic Biology (CSB), Ghent University

Prof. Dr. Serge Tavernier
Intelligence in Processes, Advanced Catalysts & Solvents (iPRACS), University of Antwerp

Dr. Lina De Smet
Laboratory of Molecular Entomology and Honeybee Pathology (L-MEB), Ghent University

This dissertation was part of the project Enzymes for Added Sustainability and Efficiency (EnzymASE). The project was funded by The Flemish Agency for Innovation and Entrepreneurship (VLAIO) and supported by Catalisti.

FACULTY OF SCIENCES

A CYP solution: an enzyme-based process for the production of long-chain α,ω-dicarboxylic acids

Ir. Delphine Devriese

Supervisor:
Prof. Dr. Bart Devreese

Academic year 2020–2021

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Doctor of Sciences: Biochemistry and Biotechnology

GHENT UNIVERSITY

LM-UGent

Ghent

Author
Ir. Delphine Devriese

Supervisor
Prof. Dr. Bart Devreese

Dankwoord

De afgelopen vier jaar kreeg ik de kans om een doctoraatsonderzoek uit te voeren in de PRU onderzoeksgroep van het labo microbiologie, voormalig het L-PROBE labo. Ik heb mezelf hier kunnen ontwikkelen als onderzoeker maar ook daarbuiten leerde ik veel bij. Dit alles was niet mogelijk geweest zonder de volgende mensen en bij deze wil ik hen ook bedanken.

Allereerst wil ik mijn promotor professor Bart Devreese bedanken. Iets meer dan vier jaar geleden botste ik op een vacature bij het toenmalige L-PROBE labo betreffende het onderzoek van enzymen voor de chemische industrie. Vanwege mijn interesse in de eiwitchemie, solliciteerde ik voor dit doctoraat en werd ik tot mijn vreugde aangenomen. Niet alleen wil ik mijn promotor bedanken om mij de kans te geven voor het uitvoeren van dit doctoraat, ook kon ik afgelopen vier jaar rekenen op een aangename samenwerking waarbij de deur ook vaak letterlijk openstond voor ons.

In het labo kon ik altijd terugvallen op mijn collega's. Stijn wil ik bedanken om mij bij de start van mijn doctoraat te begeleiden en mij te introduceren in het labo alsook in het EnzymASE project. Isabel verdient hier ook zeker een dankjewel. Als lab manager kon ik steeds op haar rekenen voor haar raad en daad met betrekking tot veiligheid, producten, toestellen en meer. Daarnaast was ze mee betrokken in EnzymASE en kon ik bij haar terecht voor het bediscussiëren van dit project. Steeds als ik iets meer wilde weten of hulp nodig had bij massa spectrometrie, kon ik aankloppen bij Sören. Voor extra informatie in verband met de membraanvesikels kon ik dan weer rekenen op Jolien Vitse. Ook mijn andere collega's wil ik graag bedanken, namelijk Ingrid, Zhoujian en Gonzalez, alsook ex-collega's Jolien Claeys, Sofie en Kasia en nieuwe collega's Amir, Amaury en Eva. Iedereen droeg en draagt bij aan een aangename sfeer in het labo, van een babbeltje tussendoor en tijdens de lunch, tot de jaarlijkse teambuilding, waardoor ik steeds met plezier naar het labo kwam.

Doorheen mijn doctoraat kon ik ook rekenen op mensen buiten het PRU labo. Zo wil ik Jordy Bauwelinck en Pieter Surmont bedanken voor het uitvoeren van de GC-MS analyses. Om mijn membraanvesikels te analyseren, kon ik gebruik maken van de NanoSight toestellen, beschikbaar in het labo experimenteel kankeronderzoek in het UZ Gent en in het labo algemene biochemie en fysische farmacie aan de faculteit farmaceutische wetenschappen. Daarbij wil ik Lien Lippens en Herlinde De Keersmaecker, respectievelijk, in het bijzonder bedanken voor de introductie van het NanoSight toestel.

Ik wil van de gelegenheid gebruik maken om mijn examen commissie te bedanken voor de kritische analyse van mijn doctoraat en de behulpzame suggesties die gegeven werden om deze dissertatie te maken tot wat ik hier nu mag afleveren.

Afgelopen vier jaar kon ik rekenen op mijn familie en vrienden en kon ik er ook steeds terecht voor een ontspannend momentje. Ik zou hierbij graag mijn ouders bedanken voor hun steun doorheen de jaren. Daarnaast verdient mijn zus Joke hier zeker een grote dankjewel. Ook zij legde een doctoraat af. Ze wist me zo steeds bij te staan met raad en daad en als er nood was aan ontspanning, stond ze alvast klaar met haar gezelschapsspelen. Als laatste had ik mijn vriend Toon willen bedanken. Doorheen de afgelopen jaren heeft hij altijd een luisterend oor geboden en klaargestaan voor mij. Hij wist me steeds te motiveren en kon mij ook tot rust brengen bij stress momenten.

Dankjewel iedereen!

Delphine

Summary

α,ω-Dicarboxylic acids (DCAs) are valuable chemical compounds and are used in a wide variety of applications. They serve as monomers for polyamides such as nylons and polyesters. Moreover, these compounds are used as precursors for pharmaceuticals, perfumes, lubricants, adhesives, plasticizers, powder coatings and corrosion inhibitors. Today, mainly short- and medium-chain DCAs are used. The application of DCAs with longer chain lengths (C > 14) is less widespread due their lower availability and elevated costs. This is the result of their challenging production process. On the one hand, long-chain α,ω-DCAs are produced through chemical conversion of fatty acids by means of metathesis. On the other hand, a fermentation process with the industrial DCA producer *Candida viswanathii* is established. Unfortunately, both processes come with a number of disadvantages and there is a need for an alternative long-chain α,ω-DCA production process for increased availability and decreased prices.

Cytochrome P450 monooxygenases (CYP) constitute a large superfamily of enzymes, performing a wide variety of reactions and accepting a diverse range of substrates. These enzymes have the ability to introduce one oxygen atom from molecular oxygen in a non-activated C-H bond, while reducing the other oxygen atom to water. This is a reaction difficult to achieve by conventional chemical routes. Therefore, CYPs are considered of great value in synthetic biology applications. Unfortunately, disadvantages such as low specific activity, the need for a redox partner and the cofactor NAD(P)H limits their use, and current application is restricted to whole-cell bioconversions.

A number of CYP families have been reported to hydroxylate fatty acid substrates at the ω-position to their corresponding ω-hydroxy fatty acids. In some cases, overoxidation to the α,ω-DCAs was observed. One of these enzymes was selected for the development of an *in vitro* biocatalytic process, producing α,ω-DCAs from fatty acid substrates, i.e. CYP52A13 from the industrial DCA producer *C. viswanathii*. An *in vitro* biocatalytic process was pursued due to its many advantages over the currently used processes. No rare metal catalysts are needed, the enzyme performs its conversion at ambient temperature in mild reaction conditions, no side reactions occur due to absence of other enzymes and no cell membrane needs to be crossed, especially important with respect to the desired long-chain DCAs. Moreover, a biocatalytic process is easier to implement in the chemical industry as fermentation requires specific know-how and specialized equipment.

In the first experimental chapter (chapter 3) it was investigated whether it would be feasible to create a soluble self-sufficient enzyme for the implementation in an enzyme reactor. CYP52A13 belongs to the class II CYP enzymes, meaning it is a membrane-bound protein, requiring the redox partner CYP reductase (CPR), which is also membrane-bound. As a soluble system was of interest due to its ease of implementation in a reactor system, both enzymes were truncated in order to remove their membrane anchor. *Pichia pastoris* was chosen as this fungal host offers many advantages towards industrial protein production. For example, little endogenous protein is secreted by this recombinant host. Producing CYP and CPR in a secreted

form was therefore pursued as this should greatly facilitate downstream processing. Furthermore, it was of interest to create a chimeric construct, where both redox partners were present in the same polypeptide chain. In this way, only one enzyme instead of two needs to be produced and fusion of the redox partners was reported to enhance coupling efficiency. For the optimization of this construct, a nontruncated variant was produced in the same recombinant host. Both truncated CYP and CPR were successfully produced using an artificial N-terminal tag, replacing the N-terminal membrane anchor. Even more, this N-tag enabled secretion to the medium for both enzymes. Unfortunately, activity could only be proven for CPR. A chimer was successfully produced but not secreted in *P. pastoris* and the chimer was retained in the microsomal membrane. However, the artificial N-tag proved to be valuable still. The chimer could be isolated in a soluble form in absence of detergents by washing the microsome with a high salt buffer. Furthermore, the N-tag improved the production yield of this class II CYP compared to the nontruncated form, which was only produced to a low level.

In the following chapter (chapter 4), an *in vivo* immobilization strategy was considered in order to enhance stability of the CYP by introduction in a membrane environment, while still retaining the advantages of an *in vitro* approach. To this end, surface display of a self-sufficient CYP52A13 on the outer membrane of *Escherichia coli* was executed. The ice nucleation protein InaK proved to be of great value as all differentially truncated constructs appeared at the surface. Unfortunately, no DCA producing catalyst could be obtained. In a next step, outer membrane vesicles were isolated from a hypervesiculating mutant. Using the natural self-sufficient CYP102A1 from *Bacillus megaterium*, spontaneously formed OMVs, as well as EDTA-extracted OMVs were isolated and showed activity against the model substrate *p*-nitrophenoxydodecanoic acid. Thereby, a proof-of-concept was delivered for *in vivo* immobilization of CYP enzymes on the surface of OMVs.

Lastly (chapter 5), the matter of cofactor regeneration was addressed. CYP52A13 needs NADPH for the delivery of electrons via CPR to CYP52A13. It would not be economically feasible to add this cofactor in stoichiometric amounts. Therefore, an enzymatic cofactor regeneration approach was investigated. To this end, formate dehydrogenase (FDH) was selected, as the substrate formate is cheap and readily available. Moreover, the reaction is irreversible and the formed CO_2 is easily removed. Because of its $NADP^+$-specificity, FDH from *Burkholderia stabilis* was selected and production in a secreted form was pursued for a facilitated downstream processing. On the one hand, FDH was produced in *P. pastoris* and secreted to the medium. FDH was purified in an active form. Unfortunately, stability issues limited the use of this enzyme and production in the recombinant host *E. coli* was explored. A secretory approach was considered by translocating the enzyme to the periplasm and enhancing leakage to the medium by a prolonged expression phase. However, this approach was not suitable for FDH production.

Although no soluble self-sufficient CYP52A13 could be delivered for industrial application, a solubilization strategy was proposed which might be useful for other class II CYPs or for other applications where soluble CYP enzymes are pursued, e.g. crystallization. Recombinant production of membrane-bound CYPs is notoriously challenging. In this dissertation, a method to immobilize CYPs *in vivo* was demonstrated, i.e. the display on the surface of *E. coli*-derived OMVs. This is believed to serve as a valuable approach for *in vitro* applications of CYP enzymes, not only as a biocatalyst, but in other CYP applications as well, such as biosensors. Additionally, this opens the possibility of *in vivo* immobilization of other difficult to produce enzymes, e.g. FDH, and the creation of a true self-sufficient system by co-expressing FDH with the self-sufficient CYP.

Samenvatting

α,ω-Dicarbonzuren (DCAs) zijn waardevolle chemische componenten en worden toegepast in een waaier aan applicaties. Ze doen dienst als monomeren voor polyamides, zoals nylon, en polyesters. Bovendien worden ze gebruikt als precursors voor farmaceutische producten, parfums, smeermiddelen, adhesieven, weekmakers, poedercoatings en corrosie inhibitoren. Vandaag worden vooral korte- en medium-keten DCAs gebruikt. DCAs met langere vetzuurketens (C > 14) worden minder algemeen gebruikt vanwege de lage beschikbaarheid en hoge kostprijs. Dit is het gevolg van hun uitdagende productieproces. Aan de ene kant worden lange-keten DCAs geproduceerd via de chemische conversie van vetzuren door het toepassen van metathese. Aan de andere kant is een fermentatieproces, gebruikmakend van de industriële DCA producent *Candida viswanathii*, beschikbaar. Jammer genoeg gaan beide processen gepaard met een aantal nadelen en er is nood aan een alternatief lange-keten DCA productieproces die leidt tot een verhoogde beschikbaarheid en verlaagde kostprijs.

Cytochroom P450 monooxygenases (CYP) vormen een grote enzym superfamilie, die een variëteit aan reacties uitvoert en diverse substraten accepteert. Deze enzymen slagen erin een zuurstofatoom van moleculair zuurstof te introduceren in een niet-geactiveerde C-H binding, waarbij het andere zuurstofatoom wordt gereduceerd tot water. Deze reactie is moeilijk te verwezenlijken met conventionele chemische methoden. Daarom worden CYPs als waardevolle enzymen beschouwd voor hun toepassing in synthetische biologie. Helaas wordt hun gebruik gelimiteerd door een lage specifieke activiteit, de nood aan een redox partner en aan de cofactor NAD(P)H. Hun huidige toepassing is beperkt tot *whole-cell* bioconversies.

Er werden een aantal CYP families beschreven die de ω-positie van vetzuur substraten kunnen hydroxyleren, wat resulteert in de vorming van hun overeenkomstig ω-hydroxyvetzuur. In sommige gevallen is er ook sprake van overoxidatie tot het overeenkomstig α,ω-DCA. Eén van deze enzymen werd geselecteerd voor de ontwikkeling van een *in vitro* biokatalytisch proces, waarbij α,ω-DCAs geproduceerd worden vanuit vetzuur substraten, namelijk CYP52A13 van de industriële producent *C. viswanathii*. Een *in vitro* biokatalytisch proces werd nagestreefd vanwege de vele voordelen ten opzichte van de huidige productieprocessen. Er zijn geen zeldzame metalen nodig als katalysator, het enzym voert de conversie uit bij een gematigde temperatuur en in milde reactiecondities, geen nevenreacties vinden plaats aangezien er geen andere enzymen aanwezig zijn en er is geen transport over een celmembraan nodig, vooral belangrijk met betrekking tot lange-keten DCAs. Daarbovenop komt het feit dat biokatalytische processen gemakkelijk te implementeren zijn in de chemische industrie. Fermentatieprocessen daarentegen vereisen specifieke *know-how* en een gespecialiseerde uitrusting.

In het eerste experimentele hoofdstuk (hoofdstuk 3) werd de mogelijkheid onderzocht om een oplosbaar en zelfvoorzienend enzym te creëren voor de implementatie in een enzymreactor. CYP52A13 behoort tot klasse II, wat betekent dat het membraangebonden is, en de redox partner CYP reductase (CPR), ook een membraangebonden eiwit, nodig heeft. Een oplosbaar systeem werd geambieerd omdat dit verdere inbouw in een reactorsysteem eenvoudiger maakt.

Daarom werden beide redox partners getrunceerd om hun N-terminale membraananker te verwijderen. *Pichia pastoris* werd geprefereerd als recombinante gastheer omwille van de vele voordelen die deze gist biedt in het licht van industriële eiwitproductie. Zo secreteert deze gist weinig endogeen eiwit. We beoogden daarom de productie van CYP en CPR in een gesecreteerde vorm om de eiwit opzuivering te faciliteren. Bovendien was er interesse in het maken van een chimeer construct, waarbij beide enzymen aanwezig zijn in dezelfde polypeptide keten. Op deze manier dient slechts één eiwit geproduceerd te worden en er wordt beschreven dat fusie van de redox partners de koppelingsefficiëntie verhoogd. Voor de optimalisatie van dit construct werd een niet-getrunceerde variant van CYP52A13 geproduceerd in dezelfde recombinante gastheer. Zowel getrunceerd CYP als CPR werden succesvol geproduceerd, waarbij een artificiële N-terminale tag het membraananker verving. Bovendien ondersteunde deze N-tag de secretie van beide eiwitten naar het medium. Helaas kon enkel voor CPR activiteit aangetoond worden. Een chimeer construct werd geproduceerd maar niet gesecreteerd door *P. pastoris* en het chimeer construct bleef gebonden aan het microsomaal membraan. Desalniettemin toonde deze artificiële N-tag zijn waarde. Het chimeer enzym kon geïsoleerd worden in een oplosbare vorm in afwezigheid van detergent, door het microsoom te wassen met een geconcentreerde zoutbuffer. Bovendien werd aangetoond dat deze N-tag leidde tot een verhoogde opbrengst in vergelijking met de niet-getrunceerde CYP variant, waarvan de opbrengst slechts laag was.

In het volgende hoofdstuk (hoofdstuk 4) werd een *in vivo* immobilisatie strategie overwogen om de stabiliteit van CYP te verhogen door de integratie in een membraanomgeving, waarbij de voordelen van een *in vitro* aanpak behouden bleven. Om deze reden werd *surface display* van een zelfvoorzienend CYP52A13 op het oppervlak van *Escherichia coli* uitgevoerd. Het *ice nucleation* eiwit InaK bleek waardevol te zijn aangezien de verschillend getrunceerde constructen allemaal op het oppervlak van de cel verschenen. Jammer genoeg werd geen DCA producerende katalysator verkregen. Een volgende stap bestond eruit om de membraanvesikels te isoleren van een hypervesikulerende mutant. Het natuurlijke zelfvoorzienend enzym CYP102A1 van *Bacillus megaterium* werd geproduceerd en zowel spontaangevormde vesikels als EDTA-geëxtraheerde vesikels werden geïsoleerd. Deze toonden activiteit ten opzichte van het modelsubstraat *p*-nitrophenoxydodecanoic acid. Hiermee werd een *proof-of-concept* geleverd voor de *in vivo* immobilisatie van CYPs op het oppervlak van membraanvesikels.

In het laatste experimentele hoofdstuk (hoofdstuk 5) werd cofactor regeneratie bekeken. CYP52A13 heeft NADPH nodig voor de aanlevering van elektronen via CPR naar CYP52A13. Economisch zou het niet haalbaar zijn indien deze cofactor in stoichiometrische hoeveelheden dient toegevoegd te worden. Daarom werd een enzymatisch cofactor regeneratiesysteem onderzocht. Hiervoor werd formaat dehydrogenase (FDH) geselecteerd, aangezien formaat goedkoop en gemakkelijk te verkrijgen is. Bovendien is de reactie irreversibel en kan het gevormde product CO_2 gemakkelijk worden verwijderd. FDH van *Burkholderia stabilis* werd gekozen omwille van zijn $NADP^+$-specificiteit en productie in gesecreteerde vorm werd nagestreefd omdat dit de eiwitopzuivering vereenvoudigt. Aan de ene kant werd FDH geproduceerd in *P. pastoris* en gesecreteerd in het medium. FDH werd opgezuiverd in een

actieve vorm. Helaas bleek de stabiliteit onvoldoende en daarom werd productie in *E. coli* bestudeerd. Productie in een gescreteerde vorm werd hier nagestreefd door translocatie naar het periplasma en een verlengde expressiefase waarbij lekkage naar het medium werd vergroot. Dit bleek echter geen geschikte aanpak voor FDH productie.

Ook al werd geen oplosbaar zelfvoorzienend CYP52A13 afgeleverd voor industriële toepassing, er werd een solubilisatie strategie gegeven die mogelijks van nut kan zijn voor andere klasse II CYPs en voor andere toepassingen waar oplosbare CYP enzymen beoogd worden, zoals kristallisatie. Recombinante productie van membraangebonden CYPs is een welbekende uitdaging. In deze dissertatie werd een methode voor de *in vivo* immobilisatie van CYPs gedemonstreerd, namelijk de display op het oppervlak van membraanvesikels, afgeleid van *E. coli*. Dit wordt aangenomen een waardevolle aanpak te zijn voor de *in vitro* toepassing van CYP enzymen, niet enkel als katalysator, maar ook in andere CYP toepassingen, zoals als biosensor. Bovendien opent dit de mogelijkheid tot *in vivo* immobilisatie van andere enzymen, moeilijk recombinant te produceren, zoals FDH, en het creëren van een volledig zelfvoorzienend systeem via co-expressie van FDH met het zelfvoorzienend CYP.

Table of contents

List of abbreviations

2OFOR	2-oxo-acid:ferredoxin oxidoreductase
5-ALA	δ-aminolevulinic acid
12-pNCA	*p*-nitrophenoxydodecanoic acid
ACN	acetonitrile
ADH	alcohol dehydrogenase
AIDA	adhesin involved in diffuse adherence
AOX	alcohol oxidase
AUC	area under the curve
BM3	CYP102A1 from *Bacillus megaterium*
BMGY	buffered glycerol-complex medium
BMMY	buffered methanol-complex medium
BMR	reductase domain of CYP102A1 from *Bacillus megaterium*
BSA	bovine serum albumin
BSTFA	N,O-bis(trimethylsilyl) trifluoroacetamide
CAT	carnitine acetyltransferase
Cb	carbenicillin
CoA	coenzyme A
CPEC	circular polymerase extension cloning
CPR	NADPH-dependent cytochrome P450 reductase
CYP	cytochrome P450 monooxygenase
DAS2	dihydroxyacetone synthase variant 2
DCA	dicarboxylic acid
DDDA	dodecanedioic acid
DMSO	dimethyl sulfoxide
DOC	deoxycholate
dOMV	detergent extracted outer membrane vesicle
DTBPMB	bis(ditertiarybutyl-phosphinomethyl)benzene
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EnzymASE	enzymes for added sustainability and efficiency
eOMV	extracted outer membrane vesicle
ER	endoplasmic reticulum
ESI	electrospray ionization
FA	fatty acid
FAD	flavin adenine dinucleotide
FALDH	fatty acid aldehyde dehydrogenase
FAME	fatty acid methyl ester
FAO	fatty acid alcohol oxidase
FDH	formate dehydrogenase
FdR	ferredoxin reductase
Fdx	ferredoxin

Fldx	flavodoxin
FMN	flavin mononucleotide
G6P	glucose-6-phosphate
G6PDH	glucose-6-phosphate dehydrogenase
GC	gas chromatography
GRAS	generally regarded as safe
HRP	horseradish peroxidase
IAA	iodoacetamide
IB	inclusion body
IM	inner membrane
IMAC	immobilized metal affinity chromatography
INP	ice-nucleation protein
IPTG	isopropyl-β-D-thiogalactoside
Kan	kanamycin
LB	Luria-Burtani
LC	liquid chromatography
Lpp	lipoprotein
LPS	lipopolysaccharides
LSLB	low-salt Luria-Bertani
MALDI-TOF	matrix-assisted laser desorption ionization time-of-flight
MRM	multiple reaction monitoring
MS	mass spectrometry
MW	molecular weight
NADH	reduced nicotinamide adenine dinucleotide
NADPH	reduced nicotinamide adenine dinucleotide phosphate
NEB	New England Biolabs
NMR	nuclear magnetic resonance
NTA	nanoparticle tracking analysis
OHFA	hydroxy fatty acid
OIMV	outer inner membrane vesicles
OM	outer membrane
Omp	outer membrane protein
OMV	outer membrane vesicle
P450RhF	CYP116B2 from *Rhodococcus* sp. NCIMB 9784
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline, including 0.1 % (v/v) Tween-20
PET	polyethylene terephthalate
PFOR	phthalate family oxygenase reductase
PG	peptidoglycan
PHA	polyhydroxyalkanoate
PHB	polyhydroxybutyrate
pNCA	*p*-nitrophenoxycarboxylic acid
PP	periplasm

PQS	*Pseudomonas* quinolone signal
PTDH	phosphite dehydrogenase
ROS	reactive oxygen species
RT	retention time
SAXS	small angle X-ray scattering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
sOMV	spontaneous outer membrane vesicle
SP	signal peptide
SRS	substrate recognition site
TB	Terrific Broth
TCA	trichloroacetic acid
TCP	total cell protein
TMCS	trimethylchlorosilane
VLAIO	Vlaams agentschap innoveren en ondernemen
WT	wild type
YNB	yeast nitrogen base
YPD	yeast extract peptone dextrose

Chapter 1
Literature overview

1 Long-chain α,ω-dicarboxylic acids

1.1 Introduction and application

α,ω-Dicarboxylic acids (DCAs) are important intermediates with a wide variety of applications, which we encounter on a daily basis. We find them in our clothes, our cars, sport equipment, packaging, etc. α,ω-DCAs are monomers included in nylons and other polyamides and polyesters. Additionally, they serve as precursors for pharmaceuticals, perfumes, lubricants, adhesives, plasticizers, powder coatings and corrosion inhibitors [1]–[3]. Adipic acid for example is produced on mass for its use in nylon-6,6 (Figure 1.1) [4]. Azelaic acid (Figure 1.1A) is known to have antibiotic properties with applications in the treatment of acne and its lithium and aluminium salts are used in lubricants. Sebacic acid (Figure 1.1A) acts as monomer in nylon-6,10 and its dibutyl ester is a plasticizer in food packaging [5]. Using dodecanedioic acid (DDDA) (Figure 1.1A) in polyamide synthesis leads to nylon-6,12. DDDA also finds other applications such as acting as a curing agent and stabilizer in acrylic powder coatings [4]. Another example is brassylic acid (Figure 1.1A), which serves as a precursor for synthetic musk [6].

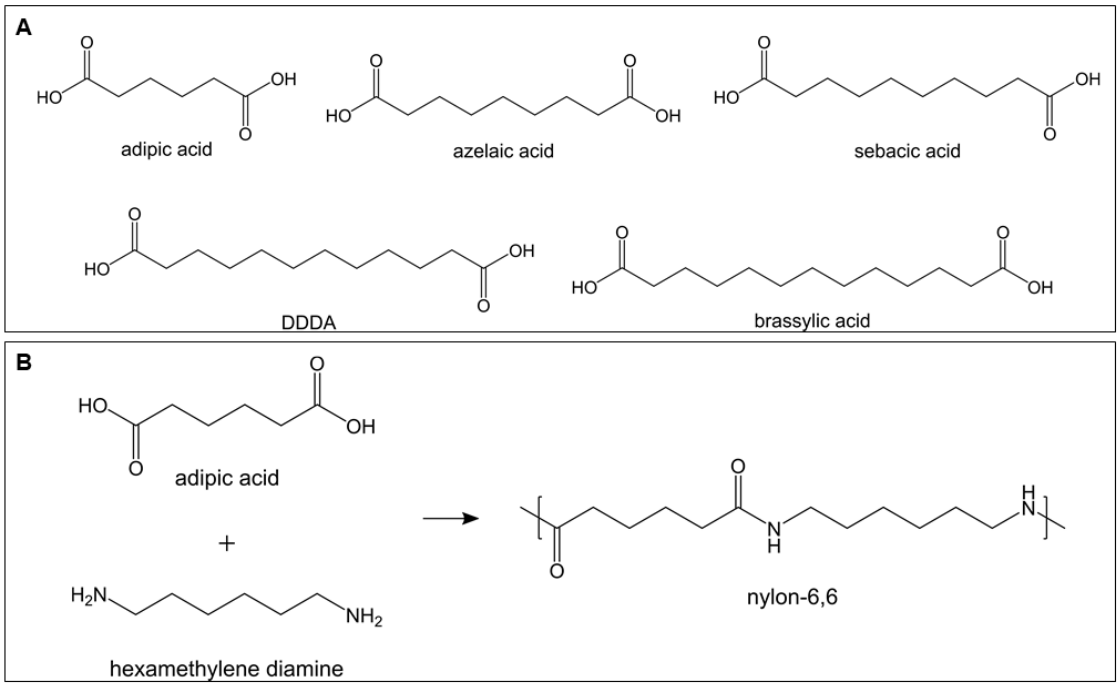

Figure 1.1: The application of α,ω-dicarboxylic acids in industry. A: Examples of α,ω-dicarboxylic acids used today. B: Nylon-6,6 is a polymer, formed from two monomer building blocks, i.e. the DCA adipic acid and the diamine hexamethylene diamine.

The application of long-chain α,ω-DCAs (>C14) is less widespread due to the fact that the production processes are more challenging, resulting in lower availability and higher costs. However, their potential as (cross)linkers in polymer synthesis is widely recognized. Their use in polyamides leads to decreased water absorbance [7]. Water absorbance can significantly affect stability and mechanical performance. This is exemplified by nylon-6,12, which shows lower moisture uptake and improved chemical resistance compared to nylon-6,6 [2]. DCAs currently used in polyesters are mainly aromatic, for example terephtalic acid for the synthesis of polyethylene terephthalate (PET). Replacing aromatic by aliphatic DCAs results in

polyesters which are more biodegradable, this in contrast to e.g. PET. However, when polyesters are synthesized with short-chain aliphatic DCAs, melting temperatures are low, as well as their tendency to crystallize. Increasing the chain length of these aliphatic DCAs could overcome this [7]. Using chain length of 16 up to 24 carbon atoms leads to polymers with polyethylene-like properties, thus showing higher melting temperatures and a higher tendency to crystallize. In comparison to PET, long-chain aliphatic polyesters are also more easy to process. Long-chain α,ω-DCAs thus offer a lot of advantages and can additionally lead to entirely new polymers. In 2015, the demand for long-chain DCAs was more than 24 kilo tons for Asia Pacific alone and the demand for long-chain DCAs is increasing annually. According to Grand View Research, the global market of long-chain DCAs is anticipated to grow with a compound annual growth rate of 7 % and the market is predicted to reach USD 300 million by 2025. However, this growth is limited by their challenging production processes [3], [8]. From this, it is clear that there is a demand for alternative production processes leading to increased availability and decreased prices.

1.2 Production processes in industry

1.2.1 Chemical conversion

In industry, α,ω-DCAs are mainly synthesized by means of chemical conversions, starting from both fossil (petroleum and natural gas) and sustainable resources (fatty acids and their esters). The longest DCA that is commercially available by chemical synthesis from fossil resources is DDDA (Figure 1.2). [2]–[4]. DCAs with longer carbon chains have also been synthesized from fossil-based building blocks, however only on a lab scale. These reactions are inefficient as they require elaborate protection and deprotection steps, which makes them useless on a commercial scale [7]. Additionally, fossil resources will become more and more limiting and the sustainability of industrial processes gains in importance. For these reasons, more attention is given to the development of alternative production processes using renewable materials such as fats and oils.

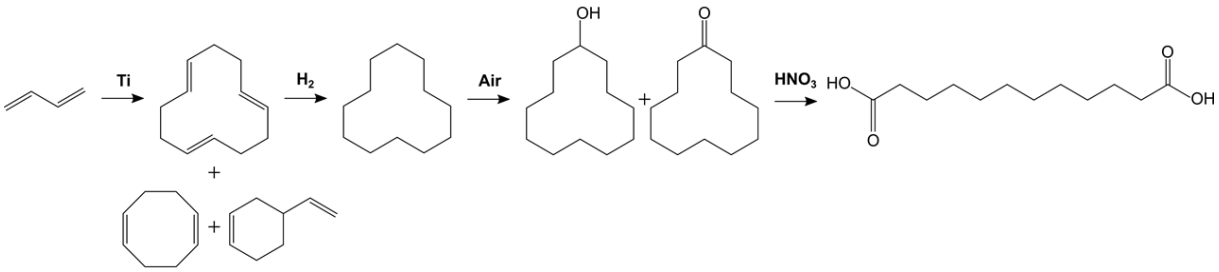

Figure 1.2: Chemical synthesis of DDDA. The reaction starts with cyclic trimerization of butadiene, which is catalysed using nickel or other metals, e.g. titanium. The obtained cyclododecatriene is then hydrogenated and sequentially oxidized with oxygen and nitric acid, respectively, to DDDA.

Fatty acids (FA) are attractive starting materials for the production of α,ω-DCAs as they already contain the desired aliphatic carbon chain and one carboxyl group. With FAs as a substrate, some α,ω-DCAs are already produced on a commercial scale. Azelaic acid and brassylic acid, for example, are synthesized via ozonolysis of oleic acid (Figure 1.3) and erucic acid, respectively, delivering several thousand tons per year [7]. Alkaline oxidation of ricinoleic acid, the main FA in castor oil, delivers around 100 000 ton sebacic acid per year [7] (Figure 1.4).

However, these conversions only provide a limited set of α,ω-DCAs. Additionally, FAs are not used to their full potential as only half of the chain is converted to the α,ω-DCA end product, the other part giving (unwanted) byproducts [2], [7].

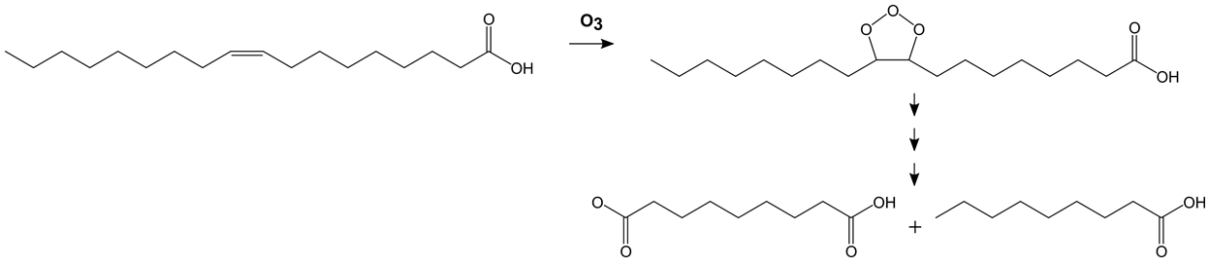

Figure 1.3: Ozonolysis of oleic acid, resulting in the α,ω-dicarboxylic acid azelaic acid and the byproduct pelargonic acid.

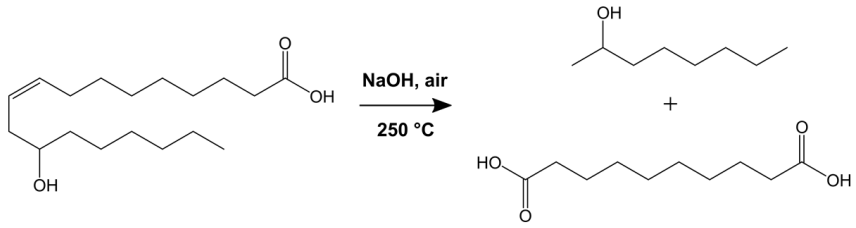

Figure 1.4: Conversion of ricinoleic acid to sebacic acid by alkaline oxidation.

A chemical conversion producing odd numbered saturated long-chain α,ω-DCA esters was reported by the group of Cole-Hamilton [9]. They accomplished this by isomerizing methoxycarbonylation of unsaturated FA esters (Figure 1.5). Here the double bond (terminally or internal) is converted to a terminal ester group using a palladium bisphosphine catalyst, and this with high conversion and selectivity. However, this selectivity is much lower when multiple double bonds are present and byproduct formation occurs in that case [7].

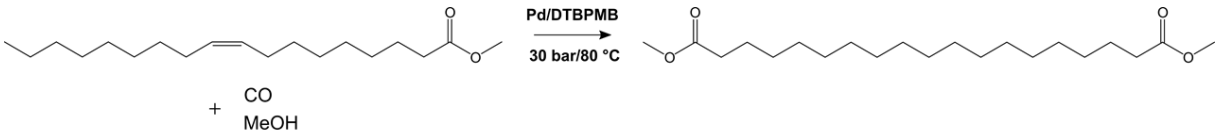

Figure 1.5: Conversion of methyl oleate to dimethyl nonadecanedioate by means of isomerizing methoxy-carbonylation using Pd complexes of bis(ditertiarybutyl-phosphinomethyl)benzene (DTBPMB).

An industrially relevant process for long-chain α,ω-DCA production is the metathesis of monounsaturated FAs. In 1972, van Dam et al reported a self-metathesis reaction using the co-catalysts tungsten hexachloride (WCl_6) and tetramethyltin (Me_4Sn) in a chlorobenzene solution for the conversion of methyl oleate to dimethyl octadecenedioate and octadecene [10]. This self-metathesis reaction was improved through replacement of WCl_6/Me_4Sn by a ruthenium-based catalyst, known as Grubbs catalyst, which is also more tolerant towards functional groups on the alkene chain [11], [12]. Ngo et al reported on the application of Grubbs second-generation catalyst in a solvent-free reaction (Figure 1.6), resulting in a 79 % oleic acid conversion and analytically pure 1,18-octadecenedioic acid was obtained at a yield of 71 %. They performed the self-metathesis reaction with 0.1 mol% catalyst under nitrogen atmosphere at a temperature between 45 °C and 55 °C. Maintaining the temperature within this range proved to be important for shifting the equilibrium of the reaction towards product formation. In this

temperature range, the reaction mixture is liquid and the formed product is solid, thereby precipitating out and thus shifting the reaction equilibrium towards product formation. When they tested their reaction using linoleic acid as a substrate, a complex mixture of products was obtained due to presence of multiple double bonds. Each alkene and unsaturated DCA product in its turn can serve as self-metathesis substrate, further complicating the final product mixture [13]. In a following article, Ngo et al tested their self-metathesis reaction using plant oils, such as soy oil, as a substrate. The first step entailed the hydrolysis of the oil in order to obtain free FAs and these were then subjected to olefin metathesis. Subsequently, the products were esterified and hydrocarbon, FA ester and DCA esters were separated from each other by means of fractionation distillation in case of large scale production. A conversion of 88 % and a DCA yield of 88 % was reported. However, the final DCA ester product was not pure and contained a variety of DCA esters as the FAs derived from plant oils are a mixture of oleic, linoleic, and linolenic acids in case of soy oil. This leads to a variety of self- and cross-metathesis reactions [14]. Such DCA mixtures are difficult to separate, leading to elevated costs due to the extensive purification process [15]. Moreover, FAs are not used to their full potential as only half of the chain is converted to the α,ω-DCA end product(s). Another disadvantage of self-metathesis is the problem of isomerization, where the double bond moves during the metathesis reaction, again leading to undesired byproduct formation [7], [15]. It must be noted that the ruthenium catalyst is difficult to remove from the final product [16], which is an additional disadvantage of the metathesis approach, although it has been stated that this contamination is negligible if the catalyst turnover number is sufficient and the process is optimized [17].

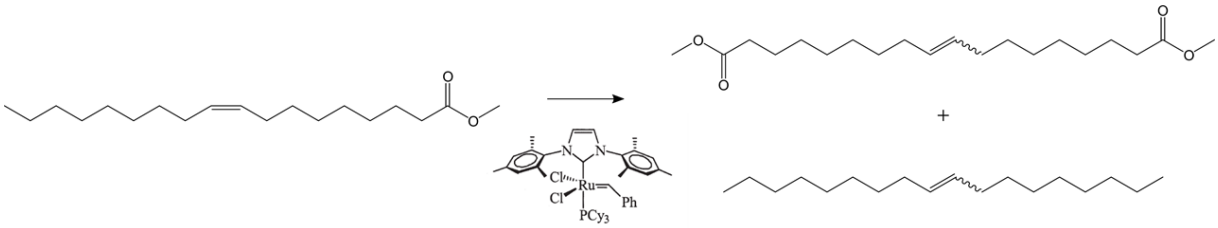

Figure 1.6: Conversion of methyl oleate to dimethyl octadecenedioate and octadecene by self-metathesis using Grubbs second-generation catalyst.

1,18-octadecanedioic acid, produced by means of metathesis, is commercially available under the name Inherent C18 Diacid from the company Elevance Renewable Science (Figure 1.7) [18]. Plant oils are first pre-treated in order to remove any contaminant interfering with the catalyst. Subsequently, 1-butene is added and cross-metathesis is carried out. Next to the cross-metathesis of 1-butene with the unsaturated fatty acid methyl esters (FAME) from the triglycerides, resulting in terminally unsaturated olefins, self-metathesis and alternative cross-metathesis reactions occur. The olefins are separated from the remaining products by means of distillation. After transesterification of the triglycerides and their derivates, monounsaturated FAMEs and DCA dimethyl esters are formed, as well as glycerol and saturated FAMEs (not shown). These are then separated. In order to obtain 1,18-octadecanedioic acid, further hydrolysis and hydrogenation is performed. This process illustrates the complexity of the metathesis process and the numerous side reactions, although in this case, every compound shown in Figure 1.7, is a valuable chemical compound [15], [19], [20].

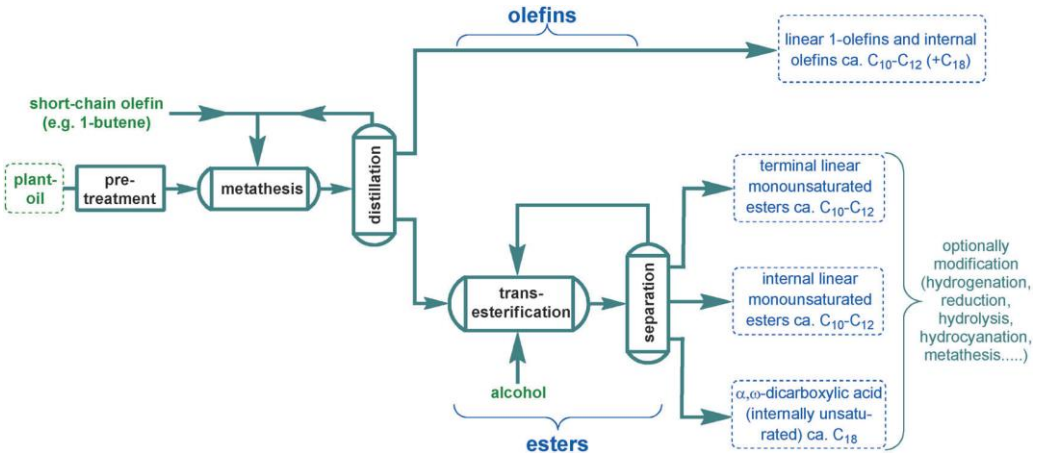

Figure 1.7: Olefin metathesis process for the production of 1,18-octadecanedioic acid, as performed by Elevance Renewable Sciences (copied from [15]).

1.2.2 Bioconversion

Several yeast strains, many of which originally assigned as *Candida* species, were found to produce and secrete long-chain α,ω-DCAs when grown either on alkanes or on FAs and their methyl esters [1], [3]. The conversion occurs with high selectivity and thus offers an interesting alternative to chemical conversion processes. Moreover, these conversions occur at ambient temperature and pressure and do not require toxic or rare-metal catalysts, giving a sustainable alternative, when using FAs as a substrate.

Industrially relevant α,ω-DCA yields are obtained with the yeast *Candida tropicalis/Candida viswanathii* (Table 1.1) [2], [3]. Consequently, this yeast has been extensively investigated and is now used for commercial α,ω-DCA production. The dual assignation is reflecting the fact that some commercially used *C. tropicalis* strains are reassigned as *C. viswanathii*. The two species are closely related and cannot be distinguished using conventional methods. Almost all literature refers to *C. tropicalis* while it is likely that some strains have to be reclassified as well [21], [22]. Below, we will discuss the two strains together as the general mechanism of DCA production is conserved and we will refer to *C. tropicalis* unless reclassification is clearly documented. Anyway, in Europe, both species are classified as pathogenic (risk group class two). For their use in a large scale fermentation process, safety precautions thus need to be taken, leading to elevated production costs. Therefore, a lot of attention is given to the search for alternative microorganisms and enzymes able to ω-oxidize FAs.

1.2.2.1 *Candida tropicalis/Candida viswanathii*

The yeast *Candida tropicalis* is known to produce and secrete long-chain α,ω-DCAs when grown either on alkanes or on FAs and has been used in industry as a valid alternative to chemically-based processes [1]–[3]. The alkanes can be used as a substrate through the presence of an ω-oxidation pathway in this yeast (Figure 1.8). Upon alkane uptake in the cell, this substrate is converted first to a terminal alcohol. In a second step, further oxidation results in the aldehyde, and thirdly, a FA is formed. These FAs can then be degraded by the β-oxidation pathway for energy. However, FAs can enter the ω-oxidation pathway as well. In this case, FAs will first be converted to ω-hydroxy FAs (OHFA). Subsequently, the corresponding aldehyde is formed and a last oxidation step gives the α,ω-DCA. DCA is then either secreted or degraded in the β-oxidation pathway [1]–[3]. The β-oxidation pathway in *C. tropicalis* is exclusively located in the peroxisomes (as is the case for other yeasts), whereas the ω-oxidation pathway is not.

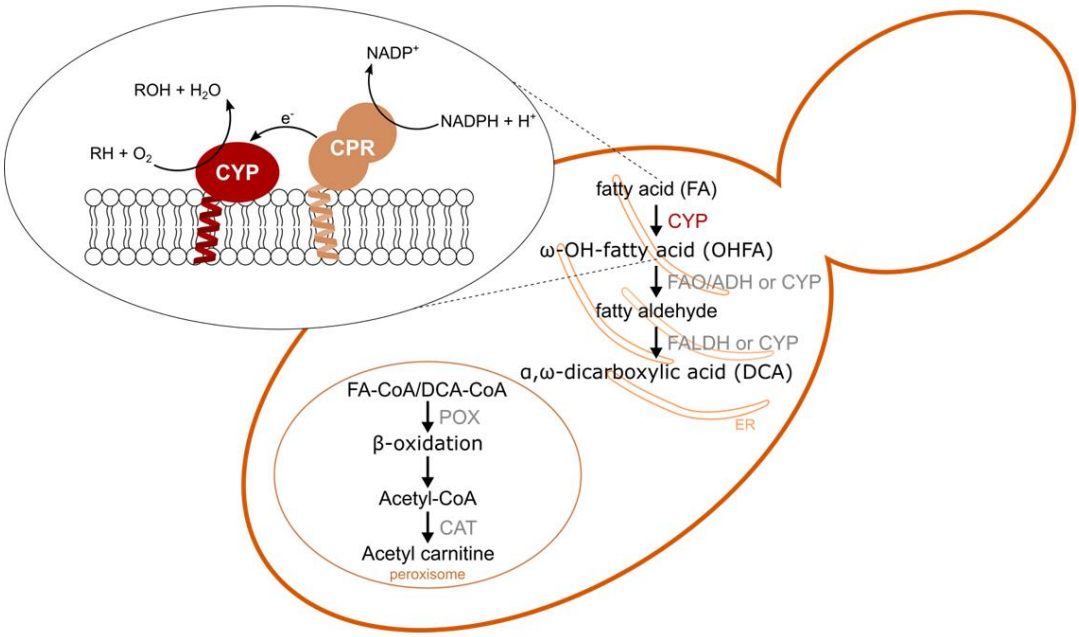

Figure 1.8: ω-oxidation of FAs and β-oxidation of FAs and DCAs in *Candida tropicalis*. The first and rate limiting step of the ω-oxidation pathway is catalysed by CYP, anchored to the ER membrane, together with its redox partner CPR.

The first step in the ω-oxidation pathway entails the terminal oxidation, either the α-position in case of alkanes or the ω-position for FAs, resulting in their corresponding alcohol. This oxidation is catalysed by the monooxygenase cytochrome P450 (CYP) together with its redox partner, NADPH-dependent cytochrome P450 reductase (CPR) (Figure 1.8). Both enzymes are anchored to the endoplasmic reticulum (ER) with an N-terminal membrane anchor, facing the cytosol. More specific, members of the CYP52[1] family have been stated to be responsible for

[1] A general nomenclature for CYP monooxygenases was proposed by Nelson et al [45] and the nomenclature and classification of CYPs is discussed in 2.2. In short, every name starts with CYP and is followed by a number which refers to the family. CYPs sharing at least 40 % amino acid sequence identity are allocated to the same family. This number is followed by a letter denoting the subfamily. Within this subfamily, enzymes share at least 55 % amino acid sequence identity. Following the letter a second number identifies the individual CYP.

this terminal oxidation. Craft et al showed that the allelic pairs CYP52A13/CYP52A14 and CYP52A17/CYP52A18 are induced upon oleic acid addition, either in a pure form or as the mixed commercial feedstream, E267 (69,9 % oleic acid) [23]. Both CYP52A13 and CYP52A17 were investigated in an accompanying paper from Eschenfeldt et al who succeeded in recombinant production of these enzymes in insect cells. CYP52A13 showed a preference for longer, unsaturated FAs. CYP52A17 had a broader substrate specificity and oxidized shorter, saturated FAs more efficiently [24]. *In vivo*, the second and third oxidation step of the ω-oxidation pathway are performed by the enzymes fatty acid alcohol oxidase (FAO) or alcohol dehydrogenase (ADH) and fatty acid aldehyde dehydrogenase (FALDH), respectively (Figure 1.8). However, Eschenfeldt et al reported that both recombinantly produced CYP52A13 and CYP52A17 were able to overoxidize FAs to the corresponding DCA, without the need for either the oxidase or one/both of the dehydrogenases [24].

DCA is not produced in high yields in wild type *C. tropicalis* strains as both FA and DCA are degraded in the β-oxidation pathway. Genetic engineering is thus a prerequisite to obtain commercially valuable DCA yields. In order to delete this degradation pathway, Picataggio et al genetically engineered the *C. tropicalis* strain ATCC 20336, now reclassified as *C. viswanathii* [21]. More specifically, the *POX4* and *POX5* genes, encoding the peroxisomal acyl-coenzyme A (CoA) oxidase 4 and 5, were deleted. These enzymes catalyse the first steps in the β-oxidation (Figure 1.8) and deletion resulted in 100 % conversion of alkane and FA substrate into DCA product [25], [26]. Picataggio et al further engineered their Δ*POX4ΔPOX5* knock-out strain (denoted as ATCC 20962) by amplifying the genes encoding the enzymes responsible for the first and rate-limiting step of the ω-oxidation pathway, i.e. CYP and CPR. This led to the strain AR40 (ATCC 20987) which showed a 30 % productivity increase and a 1,14-tetradecanedioic acid yield of 210 g/l was obtained, using the substrate methyl myristate [26]. Amplification of the gene encoding the FAO has also been put forward as a strategy to increase DCA production [27]. Other engineering strategies have been applied, next to knocking out *POX* genes. For example, Cao et al chose to knock out the carnitine acetyltransferase (CAT)-encoding gene in *C. tropicalis* strain W10-1. CAT catalyses the conversion of acetyl-CoA into acetyl-carnitine (Figure 1.8). Acetyl-CoA, the final product of the β-oxidation pathway, can be converted to acetyl-carnitine for transport to the mitochondria where it enters the TCA cycle. The heterozygote Δ*cat* indeed showed an increased DCA production of 21 %. On the contrary, the homozygote Δ*cat* lost its ability to produce DCAs [28]. More recently, a systems biological approach was utilized by Mishra et al, providing a genome-scale metabolic model of *C. tropicalis* strain iCT646, which could be used for further rational design [29]. Alternatively, random mutagenesis has been applied [30]–[32].

Next to metabolic engineering, fermentation strategies were investigated in order to increase yield. The pH for example is of major importance. Generally, first a growth phase at pH 5.5-6 takes place with an easy utilizable carbon source. This is followed by a conversion phase, where the pH is increased up to around 7.5-8 and the alkane or FA is added. This stimulates excretion of the DCA formed and enhances solubility of the substrate and product [33], [34]. Oxygen supply is another important parameter. In this regard, an interesting study is published by Jiao

et al. They added hydrogen peroxide during DCA production to solve the problem concerning limiting dissolved oxygen concentrations [35]. Optimization of feeding strategies, both of the easy utilizable carbon source and the substrate, is also of value. For example, Funk et al examined feeding strategies, together with pH control, in the production of DDDA. In fact, the substrate lauric acid has antimicrobial activity. Therefore, its methyl ester was used as an alternative substrate, combined with an optimized feeding strategy and a more gradual pH shift, leading to a yield of 66 g/l DCA without byproduct formation [36]. The same problem occurs in sebacic acid production, i.e. the substrate decanoic acid is cytotoxic. To limit this cytotoxic effect, the ω-oxidation is first induced by decane addition, before using methyl decanoate as a substrate [37]. In another study, Funk et al investigated the effect of glucose in both the growth and conversion phase of 1,18-octadecenedioic acid production using the ATCC 20962 strain. Glucose is used in the growth phase for biomass formation and inhibits ω-oxidation. The optimal glucose feeding rate was determined in order to obtain a high density without the occurrence of the Crabtree effect. As the used strain is impaired in the β-oxidation pathway, FAs cannot be used for energy and glucose needs to be added during the conversion phase for maintenance of proper cell metabolism. Interestingly, it was observed that the substrate oleic acid was stored in lipid bodies inside the yeast cell when using glucose in excess, which was not observed for lauric acid. By optimizing both the oleic acid feed rate (1 g/l/h) and the glucose feed rate (0.4 g/l/h), lipid body formation was reduced to 30 % and a volumetric productivity of 0.56 g/l/h was obtained, the highest productivity obtained with ATCC 20962 [38]. Recently, alternative ester substrates have been investigated in 1,18-octadecenedioic acid production as oleic acid leads to foaming problems. However, it was concluded that the use of oleic acid still leads to the highest yield [39]. The group of Wan focused on the effect of using different carbon sources in the growth phase in DDDA production. To achieve this goal, they isolated a new *C. viswanathii* strain ipe-1 and started by using sucrose instead of glucose as the easy utilizable carbon source [40]. They further showed that wheat straw hydrolysates could be used as a carbon source, leading to reduced production costs [41].

The fermentation process was first established by Cognis and it was by them that the strains ATCC 20962 and ATCC 20987 were developed. The fermentation was performed in two stages, as described above, i.e. a growth phase at pH 6.5 (up until the culture reached maximum turbidity), followed by a conversion phase between 7.8 and 8.9. The substrate to be converted is added at this stage. Furthermore, glucose should be added as an energy source due to the impairment of the β-oxidation pathway. The glucose addition should be low enough to prevent catabolite repression. An optimal feed rate was determined to be between 1.1 and 1.6 g/l/h. To allow for complete oxidation to the desired DCA end product, the oxygen saturation level was kept at 80 %, as a level below 50 % resulted in the formation of the intermediate ω-OHFA product. A conversion efficiency greater than 90 % was reported using methyl myristate as a substrate. No undesired modification of the substrate was observed and the purity of the final product was described to be dependent on the purity of the substrate only [42]. The high selectivity of the reaction is thus a huge advantage over the chemical metathesis process. Furthermore, the entire chain of the FA substrate is incorporated in the final DCA product, whereas only half of the FA substrate is used in case of metathesis [7].

Today, the main players in medium- and long-chain DCA production by means of fermentation are located in China, for example Cathay Industrial Biotech, Shandong Guangtong New Materials and Henan Junheng Industrial Group Biotechnology.

1.2.2.2 CYP52 family members are also found in other (dicarboxylic acid producing) fungi

Other yeasts are known to possess the ω-oxidation pathway as described above. Mainly other species of the genus *Candida* have been reported in literature. *C. maltosa* has long been identified to grow on alkanes and the major genes to be induced by n-alkanes were identified to be *ALK1*, *ALK2*, *ALK3* and *ALK5* [43]. These are now denoted as *CYP52A3*, *CYP52A5*, *CYP52A4* and *CYP52A9*, respectively, according to current nomenclature [44], [45]. Zimmer et al performed further investigation by heterologously expressing aforementioned genes in *Saccharomyces cerevisiae*. CYP52A3 proved to be the most important enzyme for the hydroxylation of n-alkanes, followed by CYP52A5, CYP52A4 and CYP52A9, in decreasing efficiency towards n-alkane hydroxylation. The order was the other way around regarding hydroxylation of FAs. Thus CYP52A9 had the highest efficiency towards FA hydroxylation, preferring the longer chain of oleic acid over shorter substrates [44]. As was the case for CYP52A13 and CYP52A17 from *C. viswanathii*, the major alkane-induced enzyme CYP52A3 showed to be able to overoxidize its substrate to the corresponding FA, without the need for the enzymes FAO, ADH or FALDH [46]. Similarly as described for *C. tropicalis*/*C. viswanathii*, metabolic engineering was applied to investigate the effect on DCA yield. Hara et al knocked out the *POX4* gene in *C. maltosa* 1098 (formerly classified as *C. tropicalis* [21]), thereby eliminating fatty acyl-CoA oxidase activity. However, they did not observe a DCA yield increase. Further experiments showed that the relation between acyl-CoA oxidase activity and DCA production is not that straightforward and other regulatory mechanisms might be involved as well [47]. In agreement with Picataggio et al, it was demonstrated that increased expression of the CYP52 enzymes leads to higher DCA production. Indeed, Kogure et al observed a higher CYP52 expression yield upon alkane induction in the DCA-hyperproducing strain, *C. maltosa* M2030, created by random mutagenesis [48].

C. cloacae has been reported to produce DCAs when grown either on alkanes [33], [49] or FAs [5] (Table 1.1). More recently, *C. guilliermondii* has been suggested as an industrial DCA producer [50]. The same engineering strategies as described for *C. tropicalis*/*C. viswanathii* have been applied for this yeast, i.e. deletion of the β-oxidation and overexpression of genes encoding the CYP and partner enzymes responsible for the first and rate-limiting step in ω-oxidation. POX1 and POX2 were identified as the major acyl-CoA oxidases involved in the β-oxidation upon oleic acid induction. However, deleting both enzymes did not result in an increased DCA yield and degradation of both oleic acid and the corresponding DCA was even accelerated. *CYP52A12* and *CYP52A15* were highly induced by growth on oleic acid. These genes were consequently selected for overexpression, together with the gene encoding the CPR redox partner. By using an autonomously replicating expression vector, Werner et al showed that only overexpressing *CYP52A12* resulted in an increased DCA yield [50]. Also more recent, *C. sorbophila* was identified as a DCA producing *Candida* species. In fermentation studies, the DDDA yield of the wild type strain was higher than was reported for the *C. viswanathii* strain

ATCC 20336, showing great promise. *C. sorbophila* produced 9.87 g/l DDDA in 144 h, starting from dodecane, whereas *C. viswanathii* ATCC 20336 produced 1.72 g/l DDDA in 120 h and this decreased to 0.61 g/l after 144 h [51]. Strain engineering was performed to delete the β-oxidation pathway, resulting in a further DCA yield increase. Starting from methyl laurate, 92.5 g/l DDDA was obtained in 126 h [52], which was higher than reported for the *C. viswanathii* strain ATCC 20962 (66 g/l DDDA [36]). However, focus was on medium-chain DCAs and longer chain FAs were not tested. The strain has been reclassified and now is known as *Wickerhamiella sorbophila* [52]. The CYP52 family has additionally been found in the pathogenic species *C. albicans*. Five *CYP52* genes were identified (*CYP52A21*, *CYP52A23*, *CYP52A22*, *CYP52A24* and *CYP52C3*). *CYP52A21* and *CYP52A23* have been successfully expressed in *E. coli* and the enzymes were solubilized from the inner membranes for further characterization. CYP52A21 showed preference for medium-chain FAs (lauric and myristic acid), whereas CYP52A23 ω-hydroxylated long-chain FAs (stearic acid and arachidic acid) more efficiently. Interestingly, by creating a chimeric construct between the N-terminal part of CYP52A21 and the C-terminal part of CYP52A23, they hypothesized that the C-terminal part is involved in the preference towards longer chains [53].

C. apicola and *Starmerella bombicola* (formerly classified as *Candida bombicola* [54]) are closely related to each other and are known to produce sophorolipids, biosurfactants consisting of an ω- or ω-1-OHFA esterified to sophorose. For *C. apicola*, CYP52E1 and CYP52E2 have been put forward to be involved in FA hydroxylation for sophorolipid synthesis [55]. CYP52E3 and CYP52N1 were identified to be involved in alkane metabolism in case of *S. bombicola*, whereas *CYP52M1* was highly induced during sophorolipid production, indicating its involvement in sophorolipid synthesis [56]. Upon further characterization, it was demonstrated that CYP52M1 hydroxylates FAs at their ω- and ω-1 positions [57]. Recently, an artificial fusion protein of CYP52M1 and CPR from *Arabidopsis thaliana* was introduced in an engineered *S. cerevisiae* strain for OHFA production. Long-chain OHFAs (C16:0 and C18:1) were successfully obtained (Table 1.1). However, both the ω and ω-1 positions were hydroxylated [58]. Interestingly, the author performed further strain engineering (disruption of the genes encoding acyl-CoA synthetases FAA1 and FAA4 and the fatty acyl-CoA oxidase POX1) in order to use glucose and not FAs as a substrate.

The yeast *Yarrowia lipolytica* has been proven of interest as well. It is an oleaginous yeast, gaining much attention in industrial applications regarding production of FAs and related compounds in the search for alternatives to petroleum-based chemistry [59]. *Y. lipolytica* has the advantage of being Generally Regarded As Safe (GRAS) and is easy to manipulate. Additionally, this yeast has already been extensively studied [59]. Moreover, it is a natural DCA producer. Twelve genes encoding enzymes of the CYP52 family were identified. Eleven CYP52s belong to the CYP52F subfamily (ALK1-10 and ALK 12) and one was assigned as CYP52S1 (ALK 11). Deleting all these genes resulted in a strain no longer able to grow using n-alkanes as the sole carbon source [60]. By expressing each individual gene in this deletion mutant, substrate specificities were determined towards either n-alkanes of various lengths or lauric acid. ALK1, 2, 9 and 10 showed to oxidize n-alkanes with ALK1, 2 and 9 preferring

longer chains whereas ALK10 oxidizes various chain lengths. ALK4, 5 and 7 showed preference for the FA lauric acid. ALK3 and 6 accepted both n-alkanes and lauric acid as a substrate, with ALK3 accepting various n-alkane chain lengths and ALK6 leaning towards longer chains. Expression of ALK8, 11 and 12 only showed trace amounts of product (ALK8 and 11) or no product at all (ALK12) [61]. Also in this case, metabolic engineering was needed in order to improve the DCA yield for industrial relevance. Indeed a higher DCA yield was obtained when the β-oxidation was impaired by deleting four of six acyl CoA oxidase encoding genes (*POX2, POX3, POX4* and *POX5*) [62]. Later, Gatter et al developed a strain where all six genes were deleted which also led to an increased DCA yield [63]. On top of this deletion, the FAO was overexpressed, involved in the second oxidation step forming the aldehyde from the alcohol. This led to additional accumulation of DCA (from 6 g/l to 11 g/l), using dodecane as a substrate [63]. Interestingly, Abghari et al used the β-oxidation impaired strain from Gatter et al for further strain engineering. On the one hand, genes were deleted in order to increase free FA accumulation. On the other hand, ALK5, CPR and FAO1 were amplified for an upregulation of the ω-oxidation. This resulted in a *Y. lipolytica* strain, able to produce long-chain DCAs starting from glycerol, and not from a FA substrate [64]. However, these engineering strategies did not yet lead to DCA yields matching these obtained by *C. tropicalis*/ *C. viswanathii* (Table 1.1). Similar as done for *C. tropicalis*, Mishra et al provided a genome-scale metabolic model of *Y. lipolytica* (iYLI647). This model delivers some additional potential engineering targets [65].

It must be noted that CYP52 family members are not exclusively found in *Candida* species and other yeasts described above. For example, CYP52s, involved in alkane and insect epicuticle degradation, have been identified in the fungi *Metarhizium robertsii* (CYP52X2) [66] and *Beauveria bassiana* (CYP52X1) [67]. Hydrocarbon-assimilating fungi such as *Aspergillus niger* and *Penicillium chrysogenum,* as well contain CYP52 enzymes [68]. However, none of these enzymes have been studied in the context of industrial DCA production and are therefore not discussed.

In conclusion, there are many alkane-assimilating fungi, some of which have been thoroughly investigated as a biocatalyst for DCA production. The CYP52 family has been proven to catalyse the first and rate limiting step of the ω-oxidation pathway, present in alkane-assimilating fungi. Some CYP52s hydroxylate n-alkanes more efficiently, others prefer ω-oxidation of FAs and overoxidation of the terminal methyl group to a carboxylic group has been reported for some CYP52s. Although many DCA producers haven been discovered and engineered by either impairing β-oxidation or overexpressing genes involved in ω-oxidation, *C. viswanathii* ATCC 20987 remains the superior biocatalyst when starting from alkanes. When using renewable resources, this holds true in case of long-chain DCAs. Looking at the medium-chain C12:0 DCA, *W. sorbophila* emerges as an interesting alternative (Table 1.1).

Table 1.1: OHFA and DCA production yields in alkane-assimilating yeasts starting from renewable resources. The highest yield found for each system is given.

Strain	Substrate	Product	Yield (g/l)	Ref.
C. tropicalis/ C. viswanathii	methyl laurate	C12:0 DCA	66	[36]
	methyl myristate	C14:0 DCA	210	[26]
	oleic acid	C18:1 DCA	100	[69]
C. cloacae	lauric acid	C12:0 DCA	10	[5]
C. guilliermondii	oleic acid	C18:1 DCA	not available	[50]
W. sorbophila	methyl laurate	C12:0 DCA	92.5	[52]
S. cerevisiae – CYP52M1	glucose	mix of 15-OHFA C16:0, 16-OHFA C16:0, 17-OHFA C18:1, 18-OHFA C18:1	0.347	[58]
Y. lipolytica	oleic sunflower oil	C18:1 DCA	23	[70]

1.2.2.3 Fatty acid ω-oxidation is not exclusive to the fungal CYP52 family

In nature, next to the fungal CYP52 family, several other CYP families have been identified possessing the ability to regioselectively oxidize FAs at their ω-position, i.e. the ER-bound mammalian CYP4 family and plant CYP76, CYP78, CYP86, CYP94 and CYP96 families and the soluble bacterial CYP153 family [71]. Also in filamentous fungi, ω-oxidizing CYPs (ER-bound), other than the CYP52 family, were identified, such as CYP539A7 and CYP655C2 from *Fusarium oxysporum* [72], and CYP630B18 from *Grosmannia clavigera* [73]. Human CYP4 enzymes ω-oxidize FAs in order to remove excess free FAs and as the first step towards DCA formation. Additionally, they are involved in the metabolism of xenobiotics, therapeutic drugs and signalling molecules and produce bioactive metabolites from arachidonic acid [74]. Therefore, the human CYP4 family, and by extension the mammalian CYP4 family, has been extensively investigated in pharmacological research [71]. The potential as a biocatalyst has not thoroughly been investigated. One example was recently published by Wernig et al. They engineered *S. cerevisiae* for *de novo* biosynthesis of hydroxyoctanoic acid. Rabbit CYP4B1, plant CYP94C1 from *A. thaliana* and fungal CYP539A7 from *F. oxysporum* were tested and it was concluded that CYP539A7 resulted in the best OHFA yield, i.e. 3 mg/l 8-hydroxyoctanoic acid. [75]. Plant ω-oxidizing CYP enzymes have a major function in the biosynthesis of cutin and suberin. OHFAs and α,ω-DCAs are monomer components of these biopolymers. Additionally, these plant CYP families were reported to be involved in plant defence, FA catabolism and reproduction [76]. CYP94A5 from *Nicotiana tabacum* was the first plant CYP reported, able to not only oxidize its FA substrate to the corresponding OHFA, but also to overoxidize this to the α,ω-DCA. However, this was only observed for the preferred substrate 9,10-epoxystearic acid [77]. A few years later, Kandel et al characterized CYP94C1 from *A. thaliana* and could show overoxidation towards the model substrate lauric acid [78]. CYP94C1 was recently further investigated in frame of industrial DCA production (Table 1.2). This enzyme, together with its redox partner, was introduced in an engineered *S. cerevisiae*

strain. This resulted in a strain, successfully producing DCA mainly of chain length C10 and C12. Similarly as described before, genes encoding the acyl-CoA synthetases FAA1 and FAA4 were knocked out, leading to free FA accumulation. This gave the ability of using renewable sugars as a substrate [79].

The soluble bacterial CYP153s have received a lot of attention in research for industrial production of both ω-OHFAs and α,ω-DCAs (Table 1.2). CYP153A1 from *Acinetobacter* sp. EB104 was the first bacterial n-alkane-oxidizing CYP from this family to be identified [80]. In subsequent years, more CYP153 enzymes were identified from different alkane-utilizing bacteria, such as from *Mycobacteria* sp., *Rhodococcus* sp., *Alcanivorax* sp., *Sphingomonas* sp., *Polaromonas* sp., *Marinobacter* sp. and *Gordonia* sp.[81]–[83]. Instead of using the natural host, CYP153 enzymes are recombinantly produced in *Pseudomonas putida* [81], [84] or, preferably, in *Escherichia coli* as this bacterium grows fast and to high densities and allows easy genetic manipulation [82], [83]. Additionally, *E. coli* bioconversion showed reduced side product formation over *P. putida* in limonene hydroxylation [85]. In the natural host, CYP153 requires two redox partners, i.e. a ferredoxin and ferredoxin reductase. These or alternative redox partners need to be co-expressed in *E. coli* for bioconversion. Alternatively, a redox partner is fused to CYP153 in a single polypeptide chain. The reductase domain of the self-sufficient CYP116B2 (P450RhF) from *Rhodococcus* sp. NCIMB 9784 proved to serve as a good redox partner, as exemplified for CYP153A13a from *Alcanivorax borkumensis* SK2 [86], [87]. Malca et al recombinantly produced three enzymes from the CYP153A subfamily; *Polaromonas* sp. (CYP153A$_{P.\ sp.}$), *Mycobacterium marinum* (CYP153A16) and *Marinobacter aquaeolei* (CYP153A$_{M.\ aq.}$). Their *in vitro* activity and selectivity towards C8–C20 saturated and C14–C18 monounsaturated FAs was investigated. CYP153A16 and CYP153A$_{M.\ aq.}$ proved to be more active than CYP153A$_{P.\ sp.}$. CYP153A$_{M.\ aq.}$ had a broader substrate specificity regarding saturated FAs, oxidizing also the longer C18 and C20 FAs. Testing monounsaturated FAs, CYP153A$_{M.\ aq.}$ was more active than CYP153A16. Additionally, CYP153A$_{M.\ aq.}$ showed higher regioselectivity. Thus CYP153A$_{M.\ aq.}$ was put forward as the superior enzyme [82] and gained a lot of interest. Redox fusion partners were investigated by Scheps et al. Similarly as described above, the reductase domain of the self-sufficient CYP116B2 was fused to CYP153A$_{M.\ aq.}$. However, fusion of the reductase domain of another self-sufficient enzyme, i.e. CYP102A1 from *Bacillus megaterium*, from here on called BMR, proved to be more efficient [88]. Malca et al showed that the mutant G307A, a residue located in the binding pocket, significantly increased the activity of CYP153A$_{M.\ aq.}$. This mutant was 15 times more active towards the C9 saturated FA and this was ascribed to an increased turnover number [82]. Further mutations were explored and the G307A/S233G (substrate access channel) double mutant of the fusion protein led to a further activity increase towards C8, C12 and C16 saturated FAs (2.3-,1.8-, and 1.4-fold improvements, respectively, compared to the G307A mutant of the CYP153A$_{M.aq.}$-BMR fusion) [89]. In another study, the G307A single mutant was subjected to a combination of site-directed mutagenesis and iterative saturation mutagenesis by Duan et al, resulting in a variant with four mutations, showing higher activity towards the longer chain of oleic acid (2.7-fold improvement *in vitro* and a 2.0-fold improvement *in vivo*) [90]. Additional studies explored the application of the CYP153A$_{M.\ aq.}$-BMR fusion product for the introduction

of a DCA production pathway in *E. coli* starting from FAs [91], [92] or, alternatively, from glucose [93]. For the overoxidation to DCAs, additional enzymes were co-expressed, i.e. ADH and FALDH [91]–[93]. Further metabolic engineering was required as well. Endogenous cofactor regeneration proved to be limiting and xylose reductase was therefore co-expressed to further increase DCA yields [92]. Moreover, prevention of FA degradation and improvement of the FA uptake was needed. Knocking out *fadE*, encoding acyl-CoA dehydrogenase, led to elimination of degradation by the β-oxidation pathway [91], [93]. Co-expression of the outer membrane transporter-encoding *alkL* of *Pseudomonas putida* GPo1 improved substrate uptake [88]. Deletion of *fadD*, preventing consumption of long-chain FAs, and overexpression of *fadL*, encoding an outer membrane long-chain FA transporter, was suggested to improve hydroxylated product yields of longer chain lengths. Indeed, higher yields of hydroxylated palmitic acid were obtained [94].

More recently, CYP153A35 from *Gordonia alkanivorans* showed to be a better biocatalyst than CYP153A$_{M.\ aq.}$ when using the substrate myristic or palmitic acid in whole-cell bioconversion, whereas CYP153A$_{M.\ aq.}$ performed better towards the substrate lauric acid. As described above, Jung et al used a *fadD* knock-out strain, overexpressing *fadL* [83]. Fusion to BMR was explored and even though the coupling efficiency between the reductase and the CYP proved to be higher, co-expression with two individual redox partners showed a higher product yield (1.5-fold improvement) using palmitic acid as the substrate [83]. Additionally, mutation studies were performed. Similarly as reported for CYP153A$_{M.\ aq.}$, a mutation in both the substrate access channel and binding pocket led to a higher activity. The single mutant D131S showed a 13-fold total turnover number improvement, a 17-fold catalytic efficiency improvement and a 50 % yield improvement. The yield was further increased by 42 % in the double mutant G297A/D131S [95]. This CYP153 family continuous to attract attention and screening studies are executed to identify new CYP153s, as exemplified by the very recent study, describing a new CYP153 from *Nocardia farcinica* IFM10152 [96].

CYP102A1 from *Bacillus megaterium* is a self-sufficient enzyme, i.e. the protein possesses a heme domain and a reductase domain on a single polypeptide chain. The enzyme is known to hydroxylate FAs, although not terminally but at the ω-1, ω-2 and ω-3 position [97]. The reductase domain has been used in creating recombinant self-sufficient CYPs, as described above for CYP153A$_{M.\ aq.}$ and CYP153A35. CYP102A1 displays one of the highest activities of all CYPs and has therefore been extensively studied for industrial applications [98]. In frame of ω-OHFA production, protein engineering was undertaken to alter the broad regioselectivity towards ω-hydroxylation. For example, Meinhold et al reported increased terminal regioselectivity towards alkanes. They selected eleven amino acids near the substrate using the substrate-bound crystal structure and subjected these to saturation mutagenesis. However, screening was done for short-chain alkanes and no terminal hydroxylation of FAs was seen [99]. Applying a directed evolution approach, Brühlmann et al successfully obtained a variant with higher terminal regioselectivity towards palmitic acid. Up to 74 % of the product was hydroxylated at the ω-position [100]. However, the natural ω-oxidizing CYPs still show much higher regioselectivities (>90 %).

Table 1.2: OHFA and DCA production yields of recombinant hosts producing CYPs other than the CYP52 family members with the ability to regioselectively oxidize FAs at their ω-position, starting from renewable resources. The highest yield found for each system is given.

Strain	Substrate	Product	Yield (g/l)	Ref.
S. cerevisiae – CYP94C1	glucose	mix of C10:0, C12:0, C14:0, C16:0 DCA and C16:0 OHFA	0.022	[79]
E. coli – $CYP153_{M.\ aq.}$	methyl laurate	C12:0 OHFA	4	[88]
	myristic acid	C14:0 DCA	1	[92]
	palmitic acid	C16:0 OHFA	2.4	[94]
	oleic acid	C18:1 OHFA	0.0292	[90]
E. coli – CYP153A35	palmitic acid	C16:0 OHFA	4.6	[83]

Terminal oxidation of the hydrocarbon chain is not excluded to the CYP monooxygenase superfamily. For example, non-heme alkane hydroxylases (Alk) have been reported to perform ω-hydroxylation of alkanes of different chain lengths [101]. Some of these have been investigated for their use in DCA production, e.g. recombinant production of AlkB from *Pseudomonas putida* GPo1 in *E. coli* resulted in the conversion of dodecanoic acid methyl ester to dodecanedioic acid monomethyl ester [102]. In this dissertation, interest goes to the CYP monooxygenase superfamily. Therefore, alternative enzymes were not discussed. CYPs are reviewed in more detail in the next part.

2 Cytochrome P450 monooxygenases

2.1 Introduction and application

Cytochrome P450 monooxygenases (CYPs) comprise a large superfamily of proteins and function in a wide variety of metabolic processes across all kingdoms of life. They play a role in the metabolism of carbon sources, in the biosynthesis of structural compounds, signalling molecules and secondary metabolites, and in the detoxification of xenobiotics. It has been established for a long-time that human CYPs have a central role in drug metabolism, metabolism of carcinogens and steroids. It was in studies on drug metabolism, using mammalian liver microsomes, that CYPs were first described. These enzymes were discovered as a red pigment in liver microsomes, showing an absorption band at 450 nm in its reduced CO-bound form, hence P450 [103], [104]. This so-called pigment subsequently proved to contain a heme prosthetic group [105]. This b-type heme is linked to the enzyme via a strong bond between the heme iron and the sulfur thiolate of a conserved cysteine residue [106]. Human CYPs, and by extension the mammalian CYPs in animal models, were thus of interest early on. Today, the number of human CYP-encoding genes is set at 57 based on the reference human genome and many of these have been characterized [107]. Eight of these CYPs are reported to be responsible for about three quarters of drug clearance reactions (3A4, 2C9, 2C19, 2D6, 1A2, 2E1, 1A1 and 2B6), of which CYP3A4 is responsible for nearly half of them [108], [109]. This can in part be ascribed to the broad substrate specificity of these CYPs. Of note, not for all 57 human CYP enzymes, functions have been determined. These CYPs are referred to as orphan CYPs, e.g. the orphan CYP20A1 [110].

Next to mammals, CYPs were discovered in all kingdoms of life and new sequences continue to be determined until today. In 2018, more than 300 000 *cyp* sequences were determined and this number is expected to reach one million in 2022 [111]. In plants, numerous CYPs have been identified. They function in the biosynthesis of lignin, UV protectants, pigments, hormones, defence compounds and oxygenated fatty acids for cutins and suberins. Additionally, the enzymes have roles in detoxification pathways [112], [113]. CYPs are also found in fungi. Although yeasts possess relatively few CYPs, many CYPs were identified in filamentous fungi, involved in a wide variety of metabolic pathways. Roles in the biosynthesis of spore walls, ergosterol and mycotoxins and in the degradation of toxic environmental pollutants (e.g. polycyclic aromatic hydrocarbons) have been described [114]. Bacterial CYPs show a remarkable variety in substrate specificity and catalytic functions as well. In general, bacteria do not comprise many CYPs and those are often involved in utilization of unusual carbon sources. Some bacteria, e.g. *E. coli*, even do not possess CYPs at all. On the other hand, secondary metabolite producing bacteria, e.g. *Streptomyces* species, comprise many CYPs which are involved in the biosynthesis of antibiotics and other secondary metabolites of interest to human and animal health and agriculture [115], [116].

Overall, CYPs are able to catalyse numerous reactions. They are not only able to perform hydroxylation, but also other reactions have been reported, such as epoxidation, N-oxidation, sulfoxidation and dealkylation to name a few [117]. Additionally, some unusual reactions have been found, reviewed in [118]. Next to a wide variety of reactions, a diverse range of substrates have been ascribed to the CYP monooxygenase superfamily [117]. This array of reactions and substrates is even more diversified through rational design and directed evolution approaches, where the work of Frances Arnold has been of major importance [119]. Especially the self-sufficient CYP102A1 from *B. megaterium*, has been a subject of multiple engineering studies [98], which was already discussed in 1.2.2.

CYPs introduce an oxygen atom from molecular oxygen into a non-activated C-H bond while reducing the other oxygen atom to water, and this with high regio- and stereoselectivity, which is difficult to achieve with conventional chemical routes. Moreover, they perform these reactions under mild conditions. Taken together with the wide variety of reactions and substrates, it is thus not surprising that this enzyme superfamily receives a lot of attention in biotechnological and synthetic biology applications. Unfortunately, some hurdles still need to be overcome for industrial application, such as their often low activity and stability and the need for a redox partner and the cofactor reduced nicotinamide adenine dinucleotide (NADH) or reduced nicotinamide adenine dinucleotide phosphate (NADPH) [120]. Despite these hurdles, CYPs have already found their way to industry and together with other promising applications, these have been extensively reviewed in [117], [120]–[124]. The majority of current applications are found in the pharmaceutical industry, illustrated by following examples. Due to the involvement of human CYPs in drug metabolism, Novartis developed an *E. coli* strain expressing 14 CYPs together with their redox partner CPR in order to prepare drug metabolites [125]. By introducing artificial pathways in *S. cerevisiae*, including CYPs, the steroids pregnenolone, progesterone and hydrocortisone were successfully produced in relevant yields. Duport et al were able to produce pregnenolone by co-expression of the bovine CYP11A1 with the reductase partners adrenodoxin and adrenodoxin reductase in a strain converting ergosterol to CYP11A1-substrates ergosta-5-ene-ol and ergosta-5,22-diene-ol. Introducing another expression cassette led to further conversion of pregnenolone to progesterone [126]. This strain was further engineered in order to produce hydrocortisone. To this purpose eight heterologous proteins, four of which are CYPs (CYP11A1, CYP11B1, CYP17A1 and CYP21A1), were introduced together with the deletion of genes involved in unwanted side reactions [127]. Another example is the production of taxadien-5α-ol, a precursor for the anticancer drug taxol, in *E. coli*. This was done by metabolic engineering and included the introduction of the *Taxus cuspidata* taxadien-5α-ol hydroxylase CYP725A4 [128], [129]. In the production of the cholesterol-lowering drug pravastatin, CYPs play a central role as well. Pravastatin has been produced in industry starting from compactin by biotransformation using *Streptomyces carbophilus* and it was found that P450sca (CYP105A3 according to current nomenclature), was responsible for the final hydroxylation [130]. More recently, a new approach was reported for pravastatin production. CYP105AS1 from *Amycolatopsis orientalis* was found to perform the same hydroxylation step as CYP105A3. The enzyme was fused to the reductase domain of CYP116B2 (P450RhF) and introduced in an engineered *Penicillium crysogenum* strain,

resulting in a promising pravastatin production host [131]. A very successful CYP application is found in the production of artemisinic acid, the precursor of artemisin, an antimalarial drug. Industrially relevant titers (fermentation titres of 25 g/l) were obtained using an engineered *S. cerevisiae* strain, heterologously expressing *CYP71AV1* from the artemisin producing plant *Artemisia annua* [132].

2.2 Classification of CYP

Enzymes are generally named based on the reaction they catalyse. Following the nomenclature, developed by The International Union of Biochemistry and Molecular Biology, CYPs belong to the EC class 1 enzymes, i.e. the oxidoreductases, catalysing a redox reaction. They are further categorized as monooxygenases, because they introduce one oxygen atom in their substrate (EC 1.14.X.Y) [133]. For this oxygenation, they generally require either NADH or NADPH as a cofactor [134]. Within this large superfamily of monooxygenases, a systematic nomenclature based on evolutionary relationships has been introduced [45], [135]. Take CYP52A13 as an example. Every name starts with the abbreviation CYP, indicating we are dealing with a cytochrome P450 enzyme. The number 52 refers to the family, i.e. all members within the family 52 will have at least 40 % amino acid sequence identity. The letter ‘A’ denotes the subfamily. When sequences share more than 55 % amino acid sequence identity, they will be ascribed to the same subfamily. The last number then refers to the individual enzyme.

CYPs catalyse a redox reaction, which requires the delivery of electrons from the cofactor NAD(P)H in most cases. These electrons are generally transferred by a redox partner, that transfers the electrons from NAD(P)H, one by one, to the heme. Originally, it was thought that only two classes existed; (I) bacterial/mitochondrial CYPs with ferredoxin and ferredoxin reductase partners and (II) eukaryotic microsomal CYPs with a diflavin reductase partner CPR. These remain the two major classes [136]. However, based on additional redox partners that have been discovered, the CYP monooxygenase superfamily was divided into ten classes by Hannemann et al [134]. This classification is still used today [137], [138]. An overview of the different classes with their respective redox partners can be found in Figure 1.9.

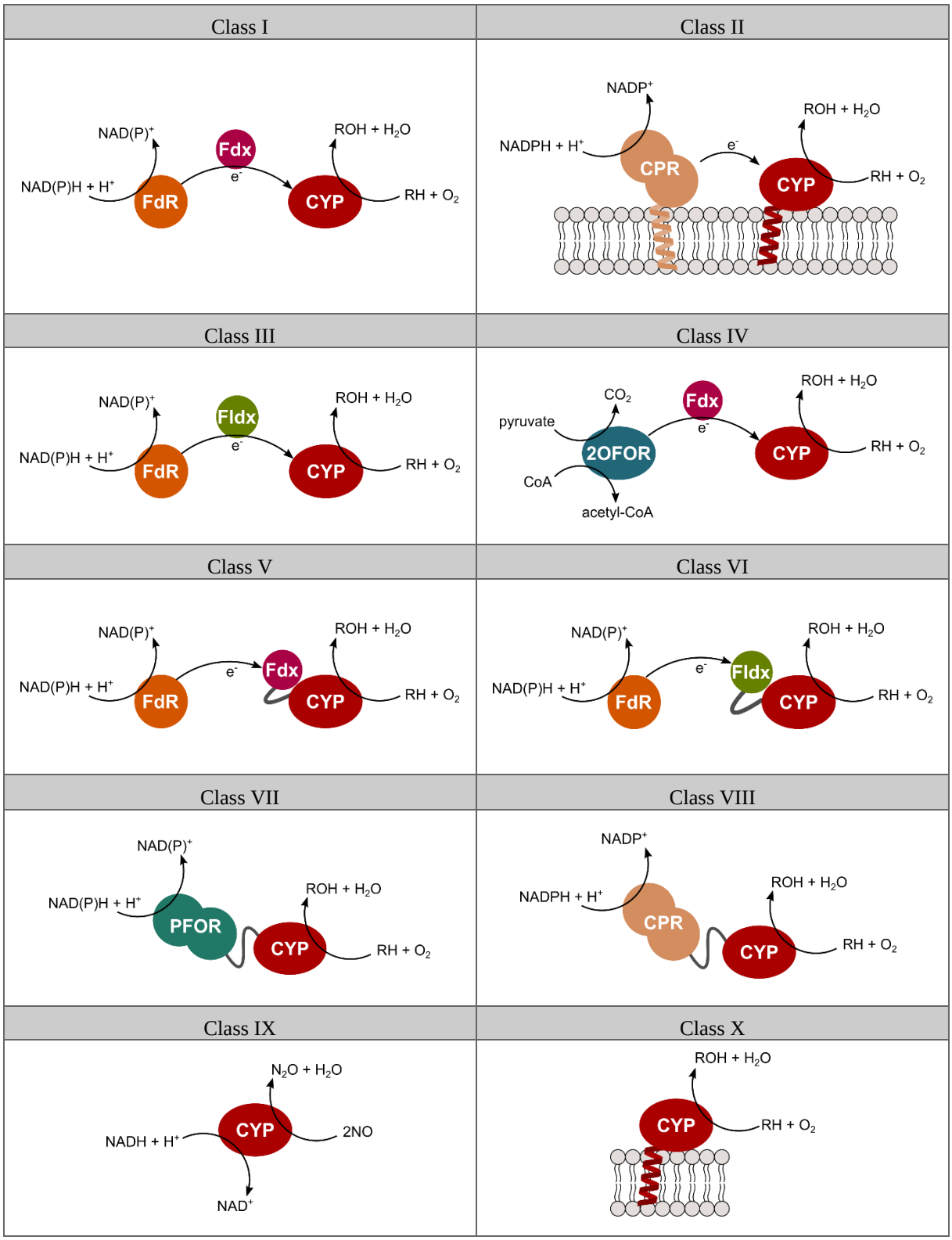

Figure 1.9: Schematic overview of the different CYP classes with their respective redox partner(s). Abbreviations: Fdx, ferredoxin; FdR, ferredoxin reductase; CPR, NADPH-dependent cytochrome P450 reductase; Fldx, flavodoxin; 2OFOR, 2-oxo-acid:ferredoxin oxidoreductase; PFOR, phthalate family oxygenase reductase.

– Class I

In the redox reaction of class I CYPs, three enzymes are involved. The first enzyme is a flavin adenine dinucleotide (FAD)-containing reductase enzyme (FdR). This reductase transfers the electrons from NAD(P)H to the second redox partner, i.e. the iron-sulfur protein ferredoxin (Fdx). The ferredoxin in its turn transfers the electrons to the CYP. This class comprises most of the bacterial CYPs, which are soluble. An example is the enzyme CYP101A1, also known as P450cam, which was the first CYP from which a crystal structure was resolved [106]. This enzyme is characterized to a large extent and has been used as a model ever since. Of note, the CYP153 family, discussed in 1.2.2, belongs to this class. Additionally, this class includes all mitochondrial CYPs from eukaryotes, which are mainly involved in biosynthesis of cholesterol-derived steroid hormones, vitamin D and bile acids. Although the same three enzyme components are required, there are differences between the mitochondrial and bacterial complex. In the bacterial complex, all three proteins are soluble, whereas in the mitochondria, only ferredoxin is soluble. The two other components are either associated with (FdR) or attached to (CYP) the inner membrane.

– Class II

The class II type CYPs need only a single redox protein partner. The redox partner is a NADPH-dependent cytochrome P450 reductase (CPR), containing two cofactors. The first cofactor is FAD, which transfers the electrons from NADPH to the second cofactor, i.e. flavin mononucleotide (FMN). FMN then transfers the electrons further to CYP. This class is the most common in eukaryotes. Both redox partners are bound with an N-terminal anchor to the ER, facing towards the cytosol. Involvement of a third component, that is cytochrome *b5*, has been reported as well [139]. Three different ways by which this third partner is engaged were observed: (1) There is an allosteric effect ascribed to cytochrome *b5* but there is no involvement in actual electron transfer. (2) The first electron is delivered by CPR. The second electron is delivered by cytochrome *b5*, which in its turn has been reduced by either CPR or the NADH-dependent cytochrome *b5* reductase. (3) CPR is in no way involved. Both electrons are transferred by cytochrome *b5* through the NADH-dependent cytochrome *b5* reductase. Of note, the CYP52 family, as well as several other ω-hydroxylating CYPs belong to this class and the CPR redox partner will be discussed in more detail later on (see 2.5).

– Class III

Similar to class I, three enzymes are involved in catalysis. However, the second enzyme, transferring electrons from the FAD-containing reductase to the CYP, is not a ferredoxin but a flavodoxin (Fldx), containing the cofactor FMN. In this respect, this class actually resembles the class II system, in that electrons are transferred from NAD(P)H, via FAD and FMN to CYP. The only difference is that the cofactors FAD and FMN are incorporated in two different proteins instead of one.

– Class IV

In this class, the redox reaction requires three enzymes as well. Remarkably, this class of CYPs does not use NAD(P)H nor a NAD(P)H-dependent flavoprotein. Instead, the electrons are transported by the enzymes ferredoxin and 2-oxo-acid:ferredoxin oxidoreductase (2OFOR). Also of interest, this class comprises thermophilic CYPs, of which CYP119A1 from the acidothermophilic *Sulfolobus solfataricus*, was the first to be discovered [140]. This enzyme has extensively been studied to be able to engineer CYPs towards increased thermostability, important for their use as a biocatalyst in industrial conditions [141].

– Class V

Up to now, this class is constituted of a single CYP enzyme, i.e. CYP51 from *Methylococcus capsulatus* [142]. CYP51 is closely related to the class I enzymes. The electron transfer goes from NAD(P)H via a NAD(P)H-dependent reductase and a ferredoxin to the CYP. The difference with class I is that the ferredoxin and the CYP are fused so instead of three separate components, this reaction requires only two enzymes.

– Class VI

This class resembles class II and class III. The electrons, delivered by NAD(P)H, are transported by enzymes containing cofactors FAD and FMN. Where in class III, FAD and FMN were part of separate enzymes, class II CYPs require one CPR, containing both cofactors in one enzyme. In this class, another type of fusion is included. Here the CYP is fused to the FMN-containing flavodoxin and a separate second enzyme contains FAD, i.e. a NAD(P)H-dependent reductase. Interestingly, characterized enzymes belonging to this class seem to be useful in bioremediation, e.g. in the degradation of the military explosive hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) [143].

– Class VII

The CYPs belonging to this class (and the next classes as well) are all self-sufficient. Class VII-type enzymes have a CYP domain at the N-terminus and a reductase domain at the C-terminus, bound by a short peptide linker. The reductase domain is related to the phthalate dioxygenase reductase family (phthalate family oxygenase reductase or PFOR) and exists of three functional parts; a FMN-binding domain, a NADH-binding domain and a ferredoxin ([2Fe-2S]) domain. It was suggested that the electrons transfer from NAD(P)H first to FMN, then to the ferredoxin and finally reach the CYP domain. In contrast to previous systems involving ferredoxins, not a FAD-containing reductase, but a FMN-containing reductase domain is involved. The first enzyme of this class to be discovered is P450RhF, or CYP116B2 according to current nomenclature, isolated from *Rhodococcus* sp. NCIMB 9784 [144].

– Class VIII

Class VIII is similar to class II, but the two enzymes are fused in one self-sufficient CYP. The reductase domain of these fusion enzymes resembles a diflavin protein, containing both a FAD and a FMN cofactor, homologous to the CPR redox partner in class II. The electron transport proceeds in the same way as in class II as well. The CYP domain is C-terminally fused to the

N-terminus of the CPR-like reductase domain by a short peptide linker. The very first fusion enzyme discovered, was CYP102A1 from *B. megaterium*, also referred to as BM3 [97]. It is a fully soluble enzyme, hydroxylating long-chain FAs at the ω-1, ω-2 and ω-3 position. This enzyme displays one of the highest activities of all CYPs. The structure of the heme domain was the second CYP structure to be resolved and the enzyme is one of the best characterized CYPs. Due to similarity in structure and electron transfer system, this fusion enzyme served as a model for the microsomal CYPs. At the time of discovery and characterization, CYP classification was restricted to class I and II, as described before. The soluble bacterial CYP102A1 enzyme showed much more similarity to the microsomal CYPs and was therefore classified as a class II enzyme. As more and more CYPs were discovered, it became clear that the two existing classes did not suffice. As such, CYP102A1 was reclassified in class VIII. However, until today this enzyme still proves to be a useful model for the eukaryotic class II type CYPs. Next to the soluble bacterial fusion enzymes, class VIII type fusions were found in lower eukaryotes as well, for example CYP505A1 from the fungus *Fusarium oxysporum* [145]. It shows high similarity to CYP102A1 regarding function and catalytic properties. However, the eukaryotic counterparts are not soluble, but bound to the to the membrane, similar to the eukaryotic CYPs from class II.

- Class IX

This class comprises fungal self-sufficient enzymes involved in denitrification. In many ways, this is an atypical class. Firstly, these enzymes are not actually monooxygenases as they reduce NO to N_2O. Secondly, they do not require a redox partner. The electrons from NADH are directly transferred to the CYP without intermediate electron transfer partners. Thirdly, these eukaryotic CYPs are soluble, localized in the mitochondria and cytosol. These are the only soluble eukaryotic CYPs identified up to now.

- Class X

A collection of independent CYPs, utilising the same catalytic features, are compiled in this last class. All enzymes perform their monooxygenation activity without the need for molecular oxygen nor NAD(P)H. Instead, their substrate serves as the oxygen donor. A reductase partner is not required either, i.e. they are self-sufficient. Belonging to this class are the plant enzymes allene oxide synthase, hydroperoxide lyase and divinyl ether synthase together with the mammalian enzymes prostacyclin synthase and thromboxane synthase. The first three CYPs are involved in the plant lipoxygenase pathway and are located in the chloroplast membranes. Both last two CYPs are enzymes of the arachidonic acid cascade.

It is clear that the CYP superfamily is large and very diverse, not only exemplified by the large variety of reactions they catalyse, but also in their required redox partners. Attention should not only be given to the CYP itself when using these enzymes in industry. The redox partner is equally important. Of specific interest in terms of industrial biocatalysis, are the self-sufficient CYPs. No additional redox partner needs to be added in the reaction and more than once, fusion constructs showed better catalytic properties [137].

2.3 Structure

In the 80's, the first crystal structure of a CYP was resolved, i.e. the structure of *Pseudomonas putida* cytochrome P450cam, or CYP101A1 [106]. Still, this is one of the best characterized CYPs [146]. Being a bacterial CYP, it is a soluble enzyme, hydroxylating the substrate camphor. Poulos et al showed that the protein fold consists of two parts, a helix-rich α-region and a helix-poor β-region. The thirteen α-helices were named in alphabetical order, A, B, B' and further from C to L, and the five β-sheets were numbered, as shown in Figure 1.10. This notation became standard for describing the P450 fold. Resolving the structure of CYP101A1, it was seen that the heme is sandwiched between helices I and L. Just before the L helix, a β-bulge is present, containing the axial ligand cysteine. Based on sequence alignment, a consensus sequence was found in the central part of helix I, which is a rather long helix close to the heme. Around this (A/G)G(D/E)T consensus sequence, halfway the I helix, a distortion is observed, resulting in a local widening which was originally proposed to be the oxygen binding site [106]. Later, it became evident that this distortion is occupied by a water molecule. Thr252 from the consensus sequence promotes protonation of oxygen leading to O-O cleavage, where it is aided by the water molecule [147], [148]. The active site is buried within the protein structure and when the structure was determined in a substrate-bound form, it was in a closed state. Therefore it was not clear how the substrate enters the active site. Also the substrate-free structure showed a closed conformation [149]. Comparing the structure of substrate-free to substrate-bound CYP101A1 showed that helices B', F and G exhibited higher thermal motion. This led to the hypothesis that these helices are involved in opening of a substrate access channel [146], [149]. This was confirmed more recently by resolving the structure of CYP101A1 in a substrate-free and open conformation. Here it was observed that the F and G helices were retracted, the F/G loop was upraised and the B' helix became disordered [150].

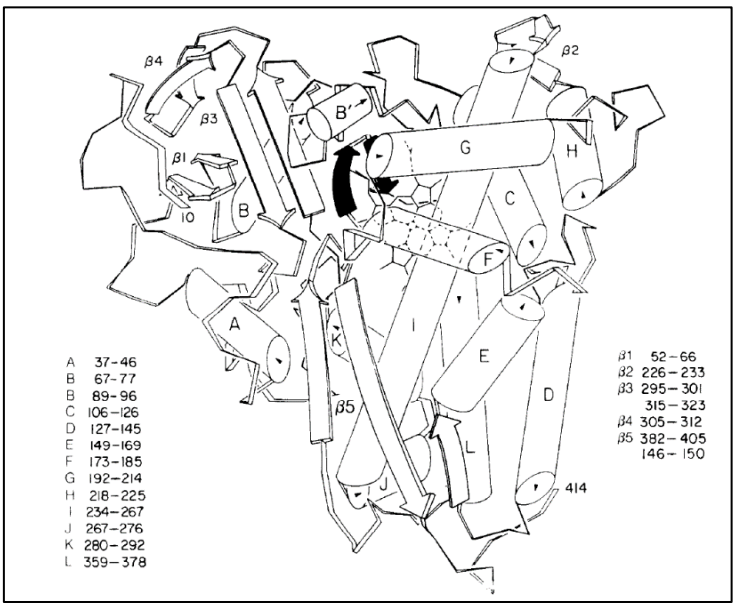

Figure 1.10: Scheme of the CYP101A1 crystal structure, illustrating the general notation used for the P450 fold. The α-helices, represented by bars, are named in alphabetical order, A, B, B' and further from C to L, and the β-sheets are numbered. Indicated in black is the antiparallel β-region and β-bulge that contains the axial heme ligand Cys357 (figure taken from [106]).

Determination of the CYP101A1 structure paved the way to a better understanding of the CYP molecular mechanism. In 1992, Gotoh applied group-to-group alignment of mammalian, more specifically the CYP2 family, versus prokaryotic CYPs, including CYP101A1. They concluded that the overall fold of CYP101A1 was maintained in the CYP2 family. However, gaps were observed, of which the gaps between helices F and G, between helices J and K and around 30 residues at the N-terminus were most significant. Based on the alignment, six substrate recognition sites (SRS) could be identified, which coincided with experimental results [151] and remain valid up to now [110] (Figure 1.11). SRS1 consist of the B' helix together with the flanking loops, which line the active site. SRS2 and SRS3 include the C-terminal end of the F helix and the N-terminal part of the G helix. Together with the F-G loop, this forms part of the substrate access channel. SRS4 is conserved as this is part of the I helix, playing an important role in the catalytic cycle. SRS5 is located in the β_3 area and SRS6 in β_5. Both β-sheets partially protrude the active site [151], [152].

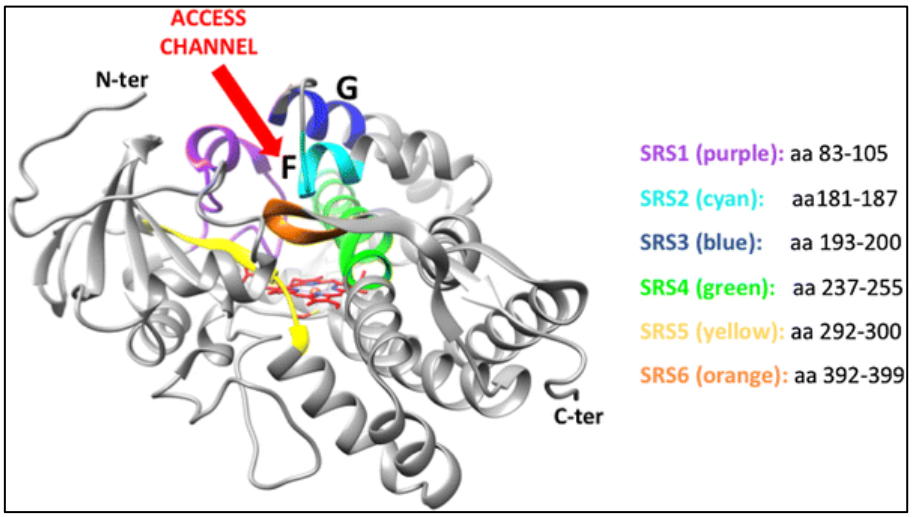

Figure 1.11: SRSs as described by Gotoh, highlighted in the structure of CYP101A1 (figure taken from [137]).

A few years later, a second CYP crystal structure was resolved, namely the structure of the *B. megaterium* CYP102A1 heme domain. Although being a soluble self-sufficient CYP of bacterial origin, it provided (and still provides) a good model for ER-bound eukaryotic CYPs (cfr. 2.2). Ravichandran et al resolved the structure in a substrate-free, open conformation [153]. They described that the heme domain consists of a helix-rich α-region and a helix-poor β-region, as was described for CYP101A1. In the substrate-free, open conformation, a long hydrophobic channel was identified as the substrate access channel. At the entrance of this channel, a hydrophobic patch was present, which was hypothesized to be the initial docking site for the hydrophobic substrates. A few years later, the crystal structure of the heme domain with the substrate palmitoleic acid bound, became available, providing more insight in the dynamics of substrate access. Comparing the open and closed structure, the largest conformational changes were in the F and G helices and in the connecting loop. The movement of this unit resulted in closure of the substrate access channel, giving rise to new intramolecular contacts [153], [154]. When comparing the structure to CYP101A1, it was seen that the overall fold was very similar (Figure 1.12). Especially around the heme group, i.e. the C-terminal part of the I-helix, the L-helix and the β-bulge containing the heme-coordinating cysteine, the

structure is highly conserved. The I-helix showed a distortion, similar to CYP101A1, to make way for a water molecule [153]. However, also structural differences were seen. Part of the variability could be ascribed to SRSs. For example, one of four major insertions was located between the helices F and G. Other differences could be ascribed to the redox partner binding site. This binding site was defined as a rectangular depression at the proximal side of the heme, with helices B, C, D, J' and K determining the outer edge (Figure 1.13). It was hypothesized that interaction with the redox partner occurs via electrostatic interactions (see 2.5) and in this region, positively charged residues protrude. A major insertion compared to CYP101A1 was present between helices J and K, i.e. the J' helix. A second insertion, also defining part of the outer edge of the redox partner binding site was found, together with some structural changes in the region of helix C and D [153]. Notably, this J' helix appeared to be present in all eukaryotic CYPs, based on sequence alignment [152].

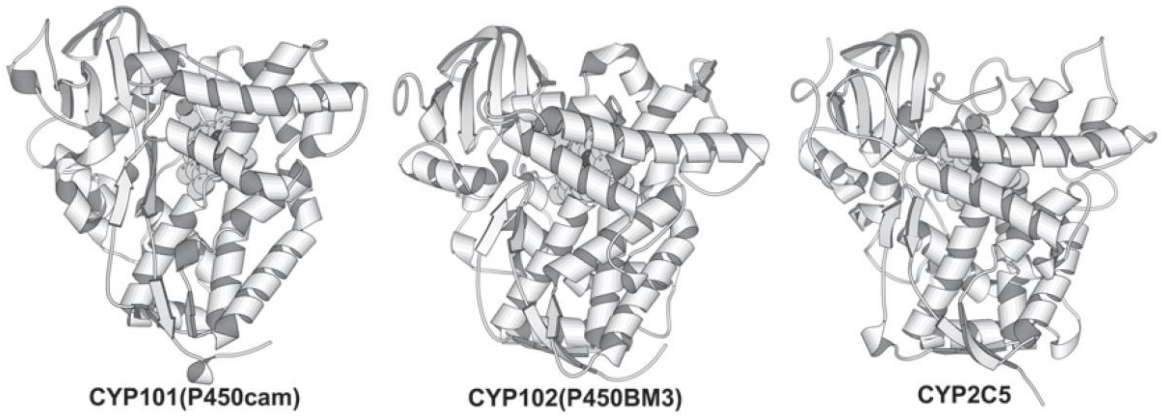

Figure 1.12: Three CYP crystal structures, described in the text, illustrating that the overall fold is conserved throughout the CYP monooxygenase superfamily (figure modified from [146]).

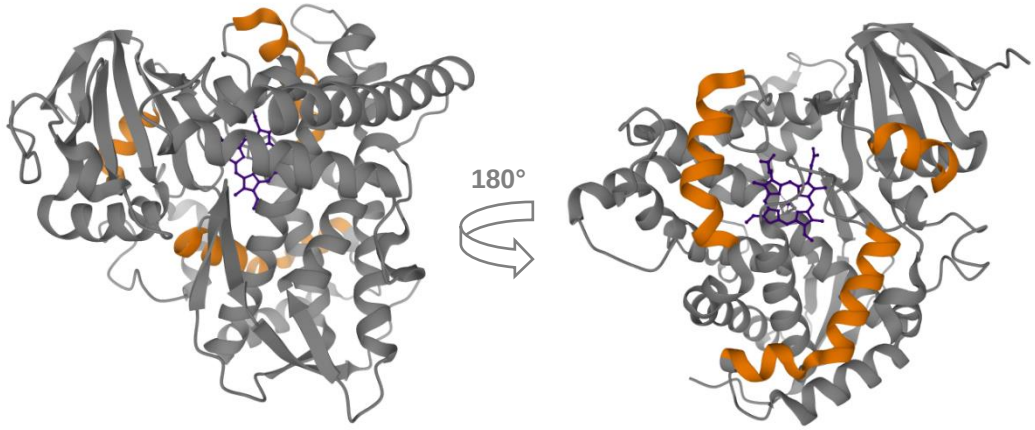

Figure 1.13: Crystal structure heme domain CYP102A1 (PDB ID 2HPD) [153]. The helices B, C, J' and K are indicated in orange and were reported to partially define the outer edge of the redox partner binding site according to Ravichandran et al [153]. (left) Distal side, in agreement with the orientation shown in Figure 1.12. (right) Proximal side, displaying the redox partner binding site. The image was created using Mol* [155].

Three other structures were published in the next couple of years, being the class I P450terp (CYP108) [156], the class I P450eryF (CYP107A1) [157] and the class IX P450nor (CYP55) [158]. It was in 2000 that the first structure of a microsomal class II CYP was published, i.e. the structure of the rabbit CYP2C5, be it in a truncated form. The N-terminal transmembrane

anchor was removed, a C-terminal 4 residue Histag was added and the solubility was increased by substituting amino acids derived from CYP2C3. This microsomal CYP showed the same protein fold as previously determined structures (Figure 1.12) [159]. The SRSs were determined by computational docking and all SRSs agreed with those determined by Gotoh [151], [159]. Compared to the heme domain of CYP102A1, all but SRS4 (helix I) were largely displaced, especially the N-terminal β-sheet and F-G loop, which leads to different substrate orientation and selectivity. The first nine residues of the truncated CYP2C5 were disordered. These residues were stated to be the linker region between the transmembrane helix and the rest of the enzyme (i.e. the catalytic domain). This linker region contains basic residues, which act to stop further translocation of the protein through the ER membrane. C-terminally a proline is present that is part of a hydrophobic patch. This patch consists of residues before and after helix A, together with β-strand 2 from β_2 and the C-terminal part of the region between helices F and G. The loop between F and G was also more hydrophobic and longer in comparison with CYP102A1 [159]. In subsequent years, more crystal structures of truncated microsomal CYPs followed. A hydrophobic patch was consistently observed, formed by the catalytic domain N-terminus together with the extended region in between helices F and G. This extended sequence gave rise to additional helices F' and G', although these are less conserved. Nevertheless these helices were not seen in prokaryotic CYPs [146], [160]. It was reported that this area is buried in the ER membrane, and this was supported by experimental data [159], [160]. It was found that epitopes, coinciding with this hydrophobic patch, were not accessible for antibody binding in CYP2B4 when bound to its native membrane, this in contrast to the solubilized form [161], [162]. A study was performed where the displacement of Langmuir–Blodgett phospholipids was measured upon CYP2B4 binding. They found the area coinciding with the displacement to be much larger than expected if only the transmembrane helix is inserted. This means that an additional region of the enzyme is embedded in the membrane [163]. The same group further explored this with atomic force microscopy. They observed that the height of CYP2B4 above the membrane was 3.5 nm. This led to the conclusion that a hydrophobic patch near the N-terminal anchor is inserted into the membrane [164]. A positively charged surface at the proximal side of rabbit CYP2C5 was stated to be orientated perpendicular to the membrane surface, assuming the hydrophobic patch is bound to the membrane, and this orientation was proposed to aid in the interaction with CPR. In this orientation, the dipole formed between the positively charged proximal side and negatively charged distal side is parallel to the membrane, enabling to guide the negatively charged CPR to the proximal side, closest to the heme [159].

It was not until 2014 that a full length microsomal CYP with a well ordered transmembrane anchor was crystallized. It was the *S. cerevisiae* lanosterol 14α-demethylase, or CYP51A1. This CYP contains an additional N-terminal helix, which is normally not present in microsomal CYPs. It was suggested that it binds to the inner leaflet of the ER with its hydrophobic face. C-terminally of this amphipathic helix, a short linker is present, followed by the transmembrane helix. With this, the secondary structure of the transmembrane domain was proven to be a helix. Its length coincided with the span of the membrane and in the C-terminal part, hydrogen bonds with the catalytic domain were formed. The helix thus appeared to maintain the orientation of

the enzyme relative to the membrane. In between helix F and G, an additional helix F' was identified, as described above. The author inferred that a portion of the F-F'-G loop was buried in the membrane, further substantiating the hypothesis of an additional membrane interaction. At the N-terminus of helix G, positive residues were observed. Based on the inferred orientation towards the membrane, it was then hypothesized that these amino acids bonded with the negatively charged phosphate head groups of the membrane phospholipid bilayer [165].

It is now generally accepted that eukaryotic CYPs of class II are partially buried in the membrane with the F/G-loop more deeply immersed, orientating the distal side towards the membrane and the proximal side towards the cytosol. It was stated that these class II CYPs were evolutionary adapted to facilitate the uptake of hydrophobic substrates, which partition in the membrane, and release the polar products at the membrane-water interface [166]. The membrane has additionally been reported to modulate catalysis in several ways. For example, the membrane composition showed to influence the redox potential, the V_{max} and K_m and the heme tilt angle [167], [168].

Today, many more CYP structures have been resolved, showing the unique P450 fold throughout. The structure around the heme is highly conserved, in agreement with the highly conserved catalytic mechanism in each and every CYP. Near the heme, several conserved regions are found, that is (A/G)G(D/E)T on the I helix and the axial ligand cysteine on the β-bulge preceding the L-helix. This cysteine residue is part of the consensus sequence $FXXGX_bXXCXG$, with X_b being a basic residue [169]. An additional conserved EXXR sequence is found on the K helix, which is proposed to aid in stabilizing the core [152]. Some regions also show high flexibility, e.g. the F/G and B/C region, denoted as the SRSs and give rise to varying sizes and shapes of the active site. These represent the high variability of substrates and reactions within this superfamily [166].

Complimentary to X-ray crystallography, nuclear magnetic resonance (NMR) and *in silico* tools such as molecular dynamic simulations, are being used to further elucidate the dynamics during catalysis and identify the different access channels. In a recent review from Urban et al [166], it was put forward how the network of channels is likely to be involved in substrate specificity, with initial substrate recognition at the surface. In this respect, the importance of the F/G region became evident as this region shows high flexibility and the F and G helices define the most common channels. Also addressing the dynamics and the different channels using *in silico* tools, is Shaik et al [170], [171]. They describe how substrate binding induces large conformational changes and triggers the whole catalytic cycle. Upon substrate binding, reduction is triggered resulting in the attachment of the reductase partner. This would then open a water channel, initially closed off by a salt bridge between a propionate of the heme and a polar basic residue near the heme. Water molecules flow into the pocket and form a water chain connecting to the acid-alcohol pair from the consensus sequence (A/G)G(D/E)T in the distortion of the I helix, crucial for further catalysis. The catalytic cycle and the redox partner interaction will be discussed in the next two parts, respectively.

2.4 Catalytic mechanism of CYP monooxygenases

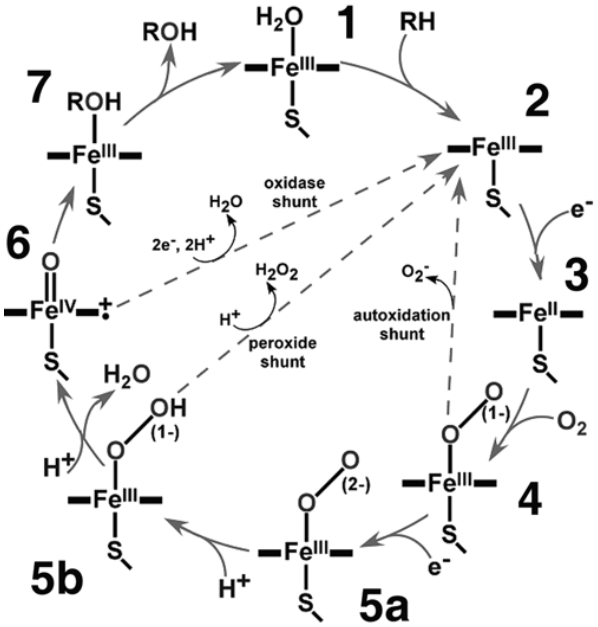

Figure 1.14: Catalytic mechanism of the CYP monooxygenase superfamily (figure taken from [172]).

The CYP monooxygenase superfamily has been proven to perform a wide variety of catalytic processes. However, the mechanism in which the oxygen is activated, is highly conserved in this superfamily. The catalytic cycle has long been investigated and there is a large consensus on what is depicted in Figure 1.14 [172], [173]. Without substrate, the enzyme is in a so-called resting state (**1**). The ferric iron of the heme group is hexacoordinated, with a water molecule being the sixth ligand. The iron is in a low spin state. When substrate binding occurs and the substrate accesses the active site, the water molecule is displaced and all water leaves the active site. The ferric iron of the heme group is now pentacoordinated and in a high spin state with an increased reduction potential (**2**). The increased potential promotes the first electron transfer from the redox partner to the heme, reducing the iron to the ferrous state (**3**). Subsequently, molecular oxygen binds to iron, which results in a ferric superoxo intermediate (**4**), and this complex is rather stable. Continuing the cycle, a second electron transfer from the redox partner occurs, creating a ferric peroxo intermediate (**5a**). This second reduction step was suggested to be the rate determining step in many CYPs. This nucleophilic ferric peroxo intermediate is short-lived and becomes quickly protonated, forming the unstable ferric hydroperoxo intermediate (**5b**). This compound is termed Compound 0. An acid-alcohol pair, either aspartic or glutamine acid combined with threonine (remember the (A/G)G(D/E)T consensus sequence), is key in this protonation and a water molecule, residing in the distortion of helix I, delivers the proton. In a next step, a second protonation occurs. Protonation of the distal oxygen results in heterolytic cleavage of the O-O bond. Water is split off and the highly reactive ferryl oxido intermediate is formed, known as Compound I (**6**). Compound I is the active oxidant. Actually, this reactive ferryl oxido species was fully characterized in heme-peroxidases long before in CYPs. In these peroxidases, the oxidant was termed Compound I and this terminology was later used for CYPs as well. Continuing in the catalytic cycle, Compound I oxygenates the substrate. It is generally accepted that the reaction proceeds according to the oxygen rebound mechanism, first reported by Groves et al [174] and more recently reviewed in [175]. Compound I subtracts

a hydrogen radical from the substrate, leaving behind a substrate radical and forming a ferryl iron hydroxy radical. Next, a rebound of the substrate radical and the hydroxy radical occurs, leading to the hydroxylated product (**7**). The end product then exits the active site and the enzyme returns to the resting state (**1**).

The catalytic cycle is not always efficient, in that three uncoupling reactions can take place [172], [173]. A first point of possible uncoupling is when the ferric superoxo intermediate is formed (Figure 1.14, **4**). Autooxidation can occur at this point, where the complex dissociates in ferric iron and a superoxide anion. The second uncoupling reaction is known as the peroxide shunt. This occurs when the proximal oxygen of Compound 0 (Figure 1.14, **5b)** is protonated. In that case, hydrogen peroxide is released and the heme returns to its resting state. A third and last possible uncoupling reaction occurs when Compound I (Figure 1.14, **6)** is reduced. Water is produced and the enzyme goes back to its resting state. This is called the oxidase shunt. It has been observed that uncoupling almost always occurs to some extent. Many factors can influence the degree of uncoupling [176], such as pH, substrate positioning in the active site and a disturbed substrate binding pocket. Also modifications regarding the redox partner can influence coupling efficiency. A notable exception is the self-sufficient CYP102A1, which appears to oxidize its substrate without loss of electrons from NADPH through one of the uncoupling reactions.

The catalytic cycle was unravelled by a variety of modelling and experimental methods. X-ray crystallography and *in silico* tools such as molecular dynamic simulations provided great insights, some of which have been described in 2.3. Also spectroscopically characterizing the intermediates played a big role and this was recently reviewed in [177]. The first intermediates were characterized early on. However, confirming the short-lived ferric iron peroxo (Figure 1.14, **5a**) and hydroperoxo (Figure 1.14, **5b**) intermediates, and especially the reactive ferryl oxido intermediate (i.e. Compound I, Figure 1.14, **6**) proved to be challenging. It was not until Davydov et al applied radiolytic reduction at very low temperatures of 77 K that spectroscopic characterization of the peroxo and hydroperoxo intermediates became possible [178]. In case of Compound I, it was not spectroscopically characterized until 2010, where the intermediate was generated by using m-chloroperbenzoic acid [179]. m-Chloroperbenzoic acid acts as an alternative oxidant and the reaction proceeds according to the reversed peroxide shunt. In this way, this uncoupling reaction can be used to the advantage of researchers.

In conclusion, the catalytic cycle has been thoroughly characterized. CYPs use a complex catalytic cycle, triggered by binding of the substrate. This results in the introduction of one oxygen atom from molecular oxygen in a non-activated C-H bond while the other oxygen atom is reduced to water. This reaction is not always 100 % efficient, as there are three possible uncoupling reactions, resulting in loss of reducing equivalents and formation of reactive oxygen species (ROS). Much is already known about the structure and catalytic cycle. However there are still questions to be answered. Is Compound I indeed the only active oxidant? How exactly is the structure related to the catalytic cycle? These and other questions were recently addressed in [180].

2.5 NADPH-dependent CYP reductase

Electrons are generally delivered to the heme of CYP through a redox partner. A variety of redox partners have been identified according to which CYPs are classified in ten classes, covered in 2.2. In this dissertation, interest goes to the class II type CYPs. Therefore, only the NADPH-dependent CYP reductase (CPR) will be discussed. Whereas the CYPs are very diverse, CPRs are not. For example, in the human ER a single CPR reduces the 50 different CYPs. Also other electron acceptors of CPR are reported, i.e. heme oxygenase, cytochrome *b5* and squalene monooxygenase, together with nonphysiological redox partners such as cytochrome c [136].

CPR contains two cofactors, namely FAD and FMN, within the same enzyme, i.e. it is a diflavin. It is thought that this enzyme is actually evolved as a fusion of two ancestral proteins. On the one hand, the N-terminal FMN-binding domain was found to be homologous with bacterial flavodoxins. On the other hand, the C-terminal domain, binding FAD and NADPH, showed homology with both ferredoxin-$NADP^+$-reductase and NADH-cytochrome *b5* reductase [181]. It is well established that the electrons first are transferred from NADPH to FAD. Then, these are transferred one by one to FMN which in its turn reduces the heme [182]. The exact electron cycle mechanism is still a matter of debate. Two electrons are delivered by NADPH in the form of a hydride ion to FAD, resulting in the formation of $FADH_2$. One electron is subsequently passed to FMN via intramolecular interaction. The air-stable semiquinone FMNH• was found to be the dominating form in mammalian microsomal CPR [183], [184]. Therefore, it was hypothesized that the CPR resting state contains the oxidized FAD and the semiquinone FMNH•, formed after so-called priming with NADPH (Figure 1.15). Upon reduction of CPR by NADPH, $FADH_2$ transfers one electron to this semiquinone, resulting in the hydroquinone $FMNH_2$. This fully reduced form is then thought to transfer one electron to the heme, going back to its semiquinone state. Subsequently, the second electron is received from FADH• for the second electron delivery to the heme. As such, CPR was proposed to operate according to a 1-3-2-1 electron cycle, with the numbers indicating the amount of electrons on the two flavins [136] (Figure 1.15). When an excess of NADPH was added, the formation of the four-electron-reduced enzyme, $FADH_2/FMNH_2$, was observed and a 2-4-3-2 electron cycle was suggested in that case (Figure 1.15). The formation of $FADH_2/FMNH_2$ proved to be thermodynamically unfavourable and took eight hours under anaerobic conditions to form [185], [186]. Therefore, this was excluded as a possible cycle *in vivo* [136], [187]. The 1-3-2-1 cycle was debated as well, as argued by Murataliev et al [187]. For example, in case of CYP102A1, the natural CYP-CPR fusion enzyme of *B. megaterium*, the semiquinone FMNH•, and not the hydroquinone $FMNH_2$, was proposed to reduce heme [188]. For CYP102A1, a 0-2-1-0 electron cycle was proposed instead [189]. It was suggested that the delayed formation of $FMNH_2$ left more time for the microsomal CPR to interact with the CYP redox partner [190].

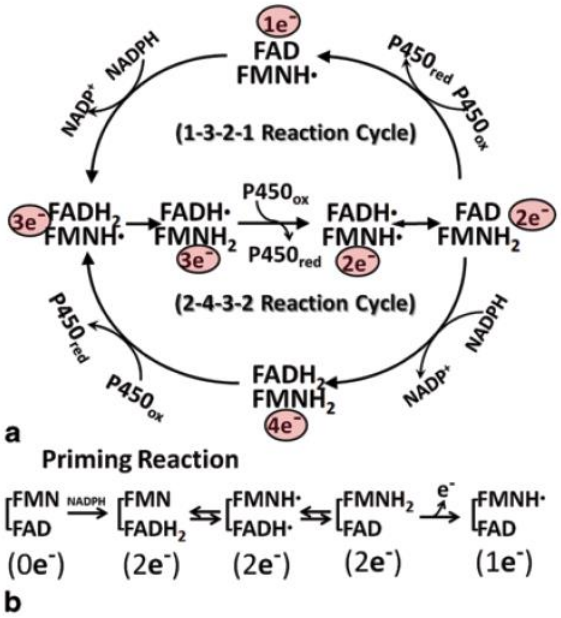

Figure 1.15: Electron cycling in CPR. a: Proposed reaction cycle mechanism. The upper half represents the 1-3-2-1 cycle, i.e. the reaction happening *in vivo*. The lower half represents the 2-4-3-2 cycle, excluded as a possible cycle occurring *in vivo*. b: Priming reaction resulting in a one electron reduced form (figure taken from [136]).

The first crystal structure of a CPR was provided by Wang et al [191] (Figure 1.16). Rat liver CPR, normally bound at the cytoplasmic side of the ER membrane with a N-terminal anchor, was recombinantly produced in *E. coli*. Afterwards, it was solubilized by limited proteolysis using trypsin. In this way, CPR was freed from its N-terminal anchor. Four domains were identified in the structure. At the N-terminus, the FMN binding domain is found. It consists of a five-stranded parallel β-sheet and at the tip of the sheet its C-terminal side, the FMN is positioned. This β-sheet is flanked by five α-helices. Two aromatic groups (Y140 and Y178) were found at each side of the FMN isoalloxazine ring. C-terminally of the FMN binding domain, a connecting domain was found, consisting mainly of α-helices. At the N-terminus of this domain, a loop forming a hinge is present. FMN is surrounded by a convex anionic surface. Interaction between the FMN binding and connecting domain appeared to be electrostatic, as the interface residues are mainly hydrophilic. The FAD binding domain follows the connecting domain in the structure. The fourth and last domain is the NADPH binding domain. When looking at the linear protein sequence, these last three domains are actually intertwined. In the crystal structure, the connecting domain is located in between the FMN and FAD binding domain and it was stated that this domain is involved in the orientation of both flavin cofactors relative to one another. The structure of the FAD binding domain consists of an anti-parallel flattened β-barrel. At the boundary between the FAD and NADPH binding domain, the isoalloxazine ring of FAD is inserted and the ring of the aromatic residue W677 was found to be stacked against the isoalloxazine ring. The rest of the FAD molecule is located closer to the connecting domain. In this so-called closed conformation, the FMN and FAD cofactor are in close contact, allowing to transfer electrons directly without intermediate amino acid residues. The shortest distance is between the 7- and 8-methyl groups of the xylene rings and these are thought to directly communicate with each other. Lastly, the NADPH binding domain shows a similar structure as the FMN binding domain, i.e. a parallel five-stranded β-sheet in between α-helices. Locating NADP(H) in the structure and positioning it relative to FAD and FMN was

difficult. The results indicated that the ribose-nicotinamide moiety of the bound $NADP^+$ was highly flexible and disordered, thus indicating a changing conformation during inter-flavin electron transfer. By modelling the $NADP^+$ molecule into the structure, Wang et al suggested that the nicotinamide ring could be positioned exactly where the W677 residue is stacked against the isoalloxazine ring of FAD [191]. Indeed, resolving the structure of three different W677 mutants, showed that upon NADPH binding, W677 must make a conformational change in order to allow stacking of the nicotinamide ring against the FAD isoalloxazine ring for hydride ion transfer. Interestingly, the FMN binding domain was disordered in one of the mutant structures. Superimposing mutant and wild type structures additionally showed a lack of order in the FMN binding domain. It was concluded that this domain was mobile relative to the FAD binding domain and conformational changes occur during catalysis [192]. Since then, other CPR crystal structures have been resolved, e.g. the CPR of *S. cerevisiae* [193], human CPR [194] and CPR of *C. tropicalis* [195], and the CPR fold proved to be well conserved [136].

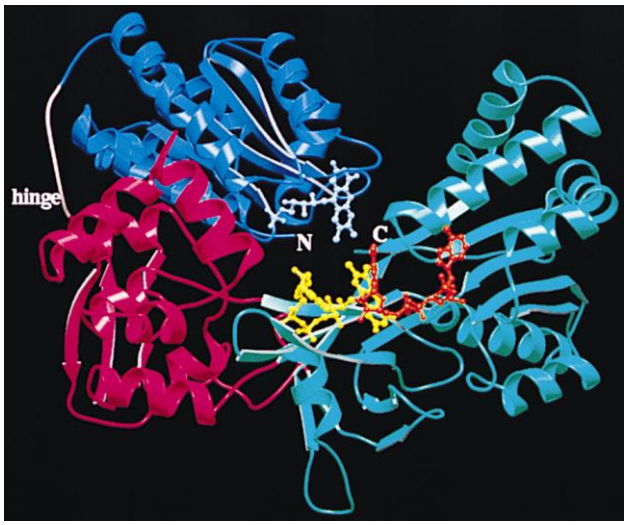

Figure 1.16: Crystal structure rat liver CPR after limited trypsin digest (figure taken from [191]).

How does this enzyme interacts with CYP? Using chemical modifications, mutations and modelling approaches, it is now well established that CYP-CPR interaction occurs through both electrostatic and hydrophobic interactions [136]. A basic region at the proximal side of CYP, mainly residues of the C helix, has been identified to interact with CPR, as was originally reviewed by Hlavica et al [196]. Cytochrome *b5* has been implicated to be involved in catalysis to a varying extend (see 2.2) and an overlapping, but not identical, region proved to be involved in cytochrome *b5* interaction. Since the review of Hlavica et al, many more studies were executed, confirming that CYP interacts with CPR with its proximal side mainly through electrostatic forces, together with hydrophobic forces [197]–[201]. On the CPR side, the anionic convex surface of the FMN domain is involved in the complex formation with CYP and studies have been done to identify the interacting residues [202]–[204]. A very recent study further suggested that not always the same binding motifs of the FMN domain are involved and that this depends on the CYP reaction partner [205]. Different approaches of modelling, mutation and modification resulted in an increased insight into CYP-CPR interaction, and X-ray crystallization of CYP-CPR complexes would be of great value to add to these insights. Unfortunately, it is very difficult to crystallize such complexes. Only a few CYP-redox partner

complex structures are available, none of which include the CPR redox partner. However, the self-sufficient CYP102A1 serves as a model again, as the crystal structure of a complex between the heme and FMN binding domain was resolved (Figure 1.17) [190]. It was seen that the α_1-helix of the FMN binding domain associates with helix C and L of the heme domain, and the outer FMN-binding loop binds to the peptide, preceding the β-bulge. In this complex, the methyl groups of the xylene ring of FMN point towards the heme on the proximal side of CYP. Additionally, it was noted that acidic amino acid residues were clustered in the FMN binding domain on the side that interacts with the proximal side of the heme domain, indicating electrostatic interaction [190]. Using the heme-FMN binding domain complex as a model, Shaik et al applied molecular dynamic simulations to investigate this interaction further. The simulation starts with the substrate entering CYP, in between the heme and helix I. This causes an enhanced interaction between helix C and I. As helix α_1 from the FMN binding domain is bound to helix C and L, the enhanced interaction causes the FMN to move closer, effectively shortening the distance between FMN and heme for electron transfer. This further illustrates how the substrate binding orchestrates the catalytic reaction [171].

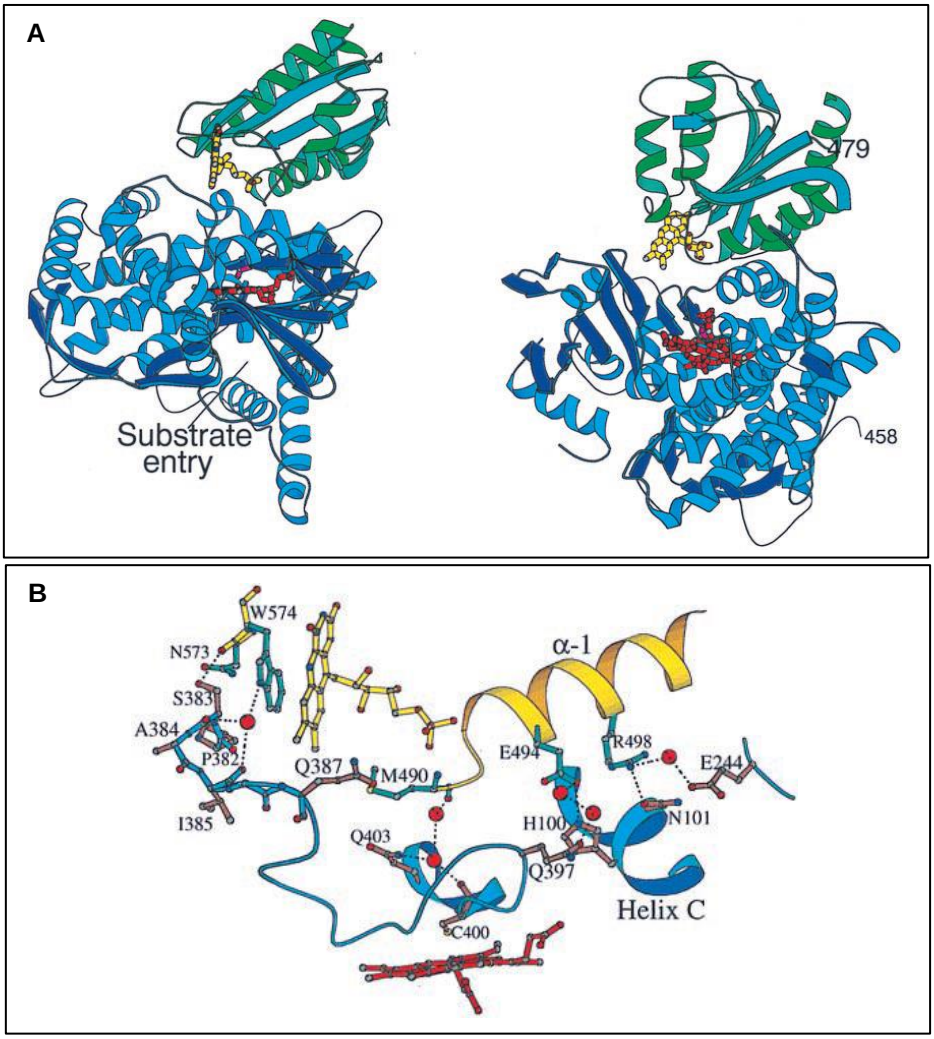

Figure 1.17: Crystal structure of the heme-FMN domain complex of cytochrome CYP102A1. A: Two views of the heme–FMN domain complex. B: The heme–FMN domain interface. (figures taken from [190])

When looking back at the resolved structure of rat liver CPR, the proposed CYP-interacting surface of the FMN binding domain is buried within the CPR enzyme and not solvent exposed, necessary for CYP interaction. Therefore, this closed structure needs to be rearranged in order to be able to interact with CYP. The connecting domain, more specifically the flexible hinge, was proposed to play a role in this rearrangement [190]. Direct evidence for this rearrangement was given by Hamdane et al [202]. The crystal structure of a mutant form of rat liver CPR was resolved, having a four amino acid deletion in the hinge. Superimposing the deletion mutant structures and the wild type structure showed that the FMN binding domain was largely displaced relative to the FAD binding domain. The FMN binding domain seemed to move away by pivoting on the C-terminal part of the hinge. A more open conformation was observed, and the two methyl groups of the xylene ring from FMN were exposed to the solvent, available for interaction with the heme group. The reducing activity of the deletion mutant was largely impaired. This reduced activity could be ascribed to the impaired electron transfer between FAD and FMN in this open conformation [202]. Around the same time, another group resolved the structure of an open conformation as well, confirming this domain rearrangement. The open structure was obtained by fusing the yeast FMN binding domain and the human connecting/FAD/NADPH binding domain [206]. NMR and small angle X-ray scattering (SAXS) experiments further confirmed the occurrence of a conformational change. An equilibrium between the open and closed structure was observed when CPR was in its oxidized form. Upon NADPH addition, it was found that equilibrium was shifted to the closed form for the cofactor-bound, reduced CPR [207]. Additional confirmation was provided by Xia et al [208]. They locked CPR in the closed conformation by engineering a disulphide bridge between the FMN and FAD binding domain. This enzyme was unable to reduce its substrates. By reduction of the disulphide bridge, the activity could be restored, proving the necessity of mobility. The general hypothesis now states that the CPR enzyme is in a closed conformation when NAPDH is bound, enabling transfer of the hydride ion to FAD and electron transfer to FMN. The enzyme then opens, leaving the reduced FMN solvent exposed for CYP interaction and subsequent reduction (Figure 1.18) [136], [168].

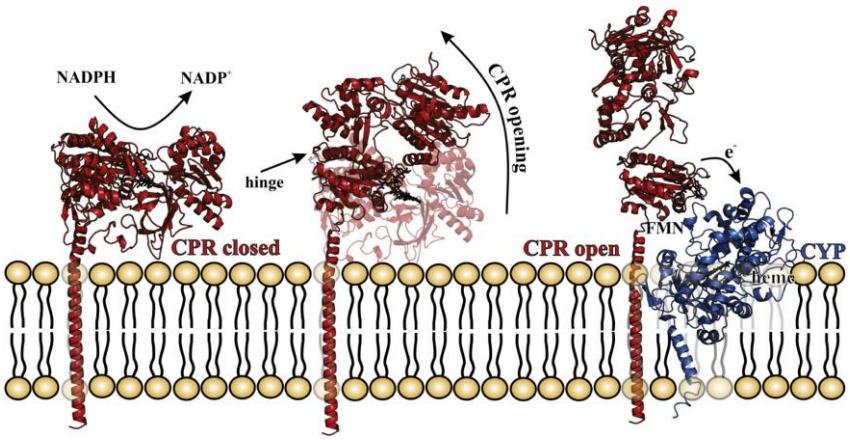

Figure 1.18: CPR conformational change during catalysis (figure taken from [168]).

CPR is bound to the membrane with an N-terminal anchor, as stated before. In rat liver CPR, a hydrophobic stretch of 22 residues, spanning the membrane, is preceded by a short nine residue hydrophilic segment. C-terminally of the hydrophobic stretch, a charged cluster was identified, followed by a flexible segment. This flexible segment is prone to proteolysis, as a tryptic digestion results in the soluble CPR, devoid of this N-terminal anchor [209]. No X-ray crystal structure of the protein including the N-terminus has been resolved yet. However, solid-state NMR studies showed the anchor to be a transmembrane α-helix [210]. Notably, the truncated rat liver CPR was unable to reduce CYP, although the artificial CPR substrate cytochrome c was reduced [191]. This would lead to the conclusion that the anchor is required for function, was it not that other truncated CPRs were able to deliver electrons to CYP, e.g. the yeast CPRs of *S. cerevisiae* [193] and *C. tropicalis* [195]. Until today, it is still not clear what the function of the membrane anchor is and how the membrane environment contributes to catalysis. Das and Sligar used Nanodiscs (discoidal lipid bilayers with a protein belt) to provide a membrane environment for the rat liver CPR. They showed how the membrane could affect the redox potential of CPR. By modifying the lipid content of the Nanodiscs to anionic lipids, it was seen that the redox potential changed. These changes resulted in some electron transfer steps to be more thermodynamically feasible [211]. Of note in this context is the difference in redox potential and electron transfer kinetics between the yeast CPR and the mammalian CPR [212]. Gideon et al suggested a passive role for the N-terminal anchor, where the anchors merely serve to bring both redox partners together in the phospholipid interface [213]. Miyamoto et al ascribed a role in increasing the coupling efficiency between the two redox partners [214]. Barnaba et al state that only the reduced CPR is an integral membrane protein. In the oxidized form, CPR is peripherally bound to the membrane to allow more diffusion in search for one of their CYP partners [215]. Xia et al reported that the transmembrane anchor is essential and the activity seen using N-terminally truncated CPR results from random collisions instead of the formation of a true CYP-CPR complex. They further hypothesize the anchor and flexible linker region to be important in the movement of the catalytic domain for interaction with their redox partner [216]. The many active chimeric constructs, fusing CYP and truncated CPR (see 2.6.4), again indicate that the anchor is not essential.

Although the CYP enzymes are the ones with the great biotechnological interest, their redox partner is equally important in order to obtain a catalytically active system. The CPR enzyme already has been investigated extensively. However, also for biotechnological applications, it is of major importance to further elucidate the interaction, e.g. to transport the electrons efficiently to CYP without loss through uncoupling.

2.6 Recombinant production of class II-type CYPs

(This part of the literature overview is in preparation for publication as a mini review)

CYPs are of major interest in synthetic biology applications as discussed in the introduction. Soluble CYPs from bacteria, as well as membrane-bound eukaryotic CYPs, many of which belong to the class II, are reported to catalyse industrially relevant reactions. The membrane-bound class II type CYPs will be the focus in this section, as these are of interest in this dissertation. CYPs are generally produced to a minor extent in their natural host. Furthermore, CYPs involved in secondary metabolism are often only produced in certain conditions. In terms of yield, making it economically interesting for industrial application, recombinant production is therefore often required, as exemplified in 1.2.2. Additionally, isolating a quantity sufficient for characterization would require a large amount of native tissue. On top, these native tissues contain not only the CYP of interest, but other CYPs as well in many cases. Extensive characterization of human CYPs and their reaction mechanisms is important in pharmacology due to their role in metabolism of drugs, carcinogens and steroids. It would not be ethical nor feasible to use liver tissue, isolated from laboratory animals and humans for each experiment regarding drug metabolism and CYP characterization. The same holds for plants. In order to characterize the CYPs involved in the production of interesting secondary metabolites, many plants would have to be collected, which again would not be feasible nor ethical in case of often rare tropical species. Therefore, CYPs are recombinantly produced to circumvent these issues. Several different host organisms have been used, depending on the purpose, e.g. employment as a biocatalyst, investigation of protein structure and/or function or providing the host with new characteristics.

2.6.1 Mammalian cells, transgenic plants and insect cells

Recombinant production of CYP enzymes in mammalian cells is of value in pharmacological research regarding drug metabolism and hepatoxicity. Human primary hepatocytes are seen as the gold standard for these experiments. However, human liver tissue is not readily available. Additionally, there is a great variability among these cells as these are isolated from human donors, having different genetic backgrounds. Moreover, they have a short life span. As an alternative, hepatoma cell lines, transfected with the CYP(s) of interest, are often used, of which the HepG2 cell line is the most widely applied [217], [218]. Other mammalian cell lines have been reported as well in this context, the very first example being the COS cell line for the expression of bovine 17α-hydroxylase CYP, now denoted as CYP17A1 [219]. CYPs are produced in their natural form and no reductase co-expression is required in these cells.

CYPs have also been recombinantly produced in plants. An example of a commercial application, is the genetically engineered blue rose. CYP75A, involved in the biosynthesis of delphinidin, an anthocyanidin giving a blue colour, was overexpressed in a rose cultivar. Together with other modifications, this resulted in a blue rose [220]. Another example is the production of rat CYP1A1 fused to yeast CPR in transgenic tobacco plant [221] and human CYP2C9 in a transgenic rice plant [222] both resulting in herbicide resistance. More recently,

the use of photosystem I, present in the chloroplast of plants, has emerged as an alternative to deliver electrons to CYP as NADPH often is a limiting factor. Plant microsomal CYPs have been shown to localize into the thylakoid membranes when the right signal sequence is provided, exemplified using CYP79A1. There, CYP efficiently obtains electrons from photosystem I via the soluble ferredoxin [223]. By fusion of CYP79A1 and ferredoxin, the activity was further increased and performed better than the reaction with its native CPR redox partner [224]. The same group used this approach to transfer a complete pathway to the chloroplast for dhurrin production [225], [226]. Of note, cyanobacteria too are able to perform photosynthesis. By metabolic engineering of the cyanobacterium *Synechocystis* sp. PCC 6803, dhurrin production was accomplished using the light-driven electron supply [227].

Insect cells, infected with baculovirus DNA, have also been applied for the recombinant production of class II CYPs, with the purpose of functionally characterizing the CYPs (e.g. CYP52A13 and CYP52A17 characterization). Again, no N-terminal truncation is necessary as these eukaryotic cells possess the required ER membrane. Recombinantly producing mammalian CYPs in insect cells results in high yields and therefore has been used rather frequently [125], [228]. Also for insect CYPs, recombinant production in insect cells is often applied. In case of plant CYPs, yeast expression is generally used, but in some cases yeast expression does not suffice and expression in insect cells led to the solution [228].

However, microbial expression systems are often preferred due to the fact that these are cheaper and faster, yields are higher and molecular cloning techniques are readily available. For example, a high amount of pure CYP is required for crystallization and functional characterization [229], [230]. Also for other applications, such as the production of drug metabolite standards [125] and the production of phytochemicals [231], microbial expression systems have proven advantageous.

2.6.2 Fungal production systems

Filamentous fungi have emerged as recombinant expression hosts in case of fungal CYPs, exemplified by Durairaj et al [114]. Especially in obtaining whole-cell bioconversion processes, filamentous fungi prove to be of value. However, bacteria and yeast are the major microbial recombinant expression hosts.

S. cerevisiae and *Pichia pastoris* (reclassified as *Komagataella phaffii*, however still referred to as *Pichia pastoris* in literature) are popular hosts for recombinant protein production because of their ease of manipulation and their ability to grow to high cell densities, this in a short time period and on inexpensive media. They possess organelles being a lower eukaryote. Presence of the ER membranes allows for recombinant CYP production in a nontruncated form [229]. Additionally, yeasts have their own set of CYP enzymes, meaning a biosynthetic pathway is well established. Primarily *S. cerevisiae* has been used as the yeast expression host for CYP enzymes. It was in *S. cerevisiae* that both the first mammalian [232] and first plant CYP [233] was successfully produced and for plant CYPs, it is still the most widely used expression host. Also *P. pastoris* is used on a regular basis and showed to be better suited for some CYPs [114],

[229]. It might thus be of interest to test different yeasts and strains, exemplified in [234], where chalcone 3-hydroxylase from *Dahlia variabilis* was produced in *S. cerevisiae* and three different strains of *P. pastoris*. They concluded that *P. pastoris* strain KM71H was the most suitable for CYP production. Another yeast expression host proven of value is *Schizosaccharomyces pombe*. As a whole-cell biocatalyst, examples are available where drug metabolites and steroids were produced by heterologously expressing human *cyps* [125]. Recently, all 57 human CYPs, together with their reductase partner, were recombinantly produced in this fission yeast, delivering a tool for more systemic investigations of the human CYPome [235].

Recombinant production of CYPs in yeast hosts has been performed both with the objective to structurally and/or functionally characterize the enzyme and to create a whole-cell biocatalyst. For example, homologous overexpression of *S. cerevisiae* lanosterol 14α-demethylase (CYP51A1) led to the first crystal structure of an N-terminally intact membrane-bound CYP enzyme (see 2.3). In order to functionally screen the plethora of fungal CYPs, recombinant production of these enzymes in yeast has proven of value as well. cDNA libraries were constructed in *S. cerevisiae* and enabled the discovery of new activities [236]. The use of yeast as a biocatalyst has already been exemplified above, e.g. the heterologous expression of genes encoding ω-oxidizing CYP52M1 and CYP94C1 (see 1.2.2) and in artemisinic acid production (see 2.1).

Although yeast possesses ER membranes, it still might be advantageous to alter the N-terminal membrane anchor for recombinant CYP production. Schoch et al demonstrated this by altering the N-terminus of the plant CYP73A1 for production in *P. pastoris* [237]. They replaced the first 20 amino acids with the peptitergent sequence PD1, which is an amphipathic helix of 25 amino acids. Together with the addition of four histidines C-terminally, they succeeded to purify CYP73A1 in a soluble form in absence of detergents, pure and stable enough for NMR experiments. In search for a microbial approach to opioid production, Galanie et al successfully introduced a complete pathway into *S. cerevisiae* [238]. In this biosynthesis, the conversion of *(R)*-reticuline to salutaridine is catalysed by salutaridine synthase, or CYP719B1. In order to enhance correct processing and improve overall activity, they tested several alternative N-terminal sequences based on the enzyme cheilanthifoline synthase, a plant CYP with 61-68 % sequence identity, showing high activity when produced in yeast. This resulted in a sixfold conversion increase.

2.6.3 *Escherichia coli*

E. coli has long been an established host for recombinant protein production, both for protein purification purposes and for their use as a whole-cell biocatalyst. Genetic engineering techniques are well known and similar to yeast, they are able to grow to high cell densities on inexpensive media. The growth kinetics of *E. coli* are unmet, with a doubling time of only 20 min in ideal conditions. As *E. coli* does not possess endogenous CYPs, no interference with heterologous CYPs occurs. However, a CYP biosynthetic pathway is not present. A heme precursor (i.e. δ-aminolevulinic acid, or 5-ALA) needs to be added in order to overcome the

limited heme biosynthesis [229], [230]. Alternatively, glutamyl-tRNA reductase can be co-expressed in order to circumvent 5-ALA addition [239]. *E. coli* is a prokaryote, meaning that it lacks ER membranes, the native location of the class II CYPs. Therefore, engineering of the N-terminus containing the transmembrane anchor, is required in most cases. The N-terminus is mutated or truncated either to direct the CYP towards the *E. coli* membrane or to completely solubilize the enzyme. Zelasko et al published a review giving a good overview of the different approaches for high-yield production of membrane-bound eukaryotic CYPs in *E. coli* and this was more recently updated by Hausjell et al [229], [230].

Barnes et al obtained the first successful production of a mammalian CYP, i.e. bovine CYP17A1, in *E. coli* [240]. To this end, they altered the second codon to alanine as this was reported to be a preferred second codon. Additionally, silent mutations were performed in codons four to seven in order to increase the AT content, with the aim of reducing secondary structure formation in the mRNA. CYP17A1 was retrieved in an active form in the membrane fraction. This sequence, consisting of the amino acids MALLLAVF (referred to as the Barnes sequence), proved to lead to successful production of many other mammalian CYPs, where the sequence was either added N-terminally or substituted after alignment of the respective CYP to CYP17A1. This approach of changing the second codon to alanine and reducing secondary structure formation in the mRNA was frequently applied [229], [230]. For plant CYPs, *E. coli* expression was achieved as well with this so-called Barnes sequence, e.g. in the metabolic engineering for 8-hydroxycadinene production [241] and for taxadien-5α-ol production [129]. Next to the Barnes sequence, other sequences have been tested for mammalian CYPs, e.g. the *ompA* and *pelB* leader sequences [242]. These sequences were added to the N-terminus of the native CYP, directing the enzyme either to the membrane or periplasm, respectively. These sequences are processed during translocation through the inner membrane, leaving behind the CYP in its native form, localized in the membrane fraction.

For plant CYPs, a screening of different tags has recently been performed in order to deliver a toolbox for production in *E. coli*. In one study, four different tags were evaluated; a 28-amino acid tag enhancing yields of difficult membrane proteins and the transmembrane domains of three highly produced *E. coli* membrane proteins (SohB, YafU and LepB). The SohB transmembrane domain provided the highest yield and activity [243]. In another study from the same group, a N-terminal tag library was constructed containing tags from *E. coli* membrane proteins, signal peptides, transporters and the previously mentioned 28-amino acid tag, delivering an expanded range of tags [244]. In case of fungal CYPs, a large-scale screening was performed by Ichinose et al in 2013 [245] and they extended this screening in 2015 [246]. Contrary to testing transmembrane domains from *E. coli*, they produced a large number of fungal CYPs and searched for those who were produced well with minor N-terminal modifications or no modifications at all. Subsequently, chimers were created substituting the original transmembrane domain with the ones identified as handled well by *E. coli* in the screening. They proposed that MSLLLAATLFLHSRQKRYPL, the N-terminal tag of CYP5144C1, could serve as an alternative to the Barnes sequence. 63 other candidate chimeric partners were given as well.

More recently, Schumacher et al reported the surface display of human CYP3A4 using Autodisplay (Figure 1.19) [247]. The so-called autotransporter AIDA-I (adhesin involved in diffuse adherence), an endogenous *E. coli* protein, was used. This protein consists of an adhesin passenger at the N-terminus and a translocator (β-barrel) domain at the C-terminus, fused by a linker peptide. The adhesin passenger can be replaced by a recombinant passenger, enabling surface display (Figure 1.19A). A signal peptide from ctxB (β-subunit of cholera toxin) was included at the N-terminus and serves to facilitate transportation across the inner membrane. The process of translocation is established and the transport of the recombinant passenger is described in Figure 1.19B [248]. The same group later showed that this Autodisplay system could be used for co-expressing the redox partner by surface displaying CYP1A2 together with human CPR [249]. Surface display of enzymes combines several advantages from both whole-cell bioconversion and *in vitro* biocatalysis. On the one hand, the enzyme is in direct contact with the extracellular space and therefore, the substrate does not need to cross any membrane, in contrast to whole-cell bioconversions. On the other hand, no purification, which is particularly challenging for membrane-bound CYPs, is required, and the CYPs are integrated in a membrane environment, beneficial for the enzyme stability [247], [249].

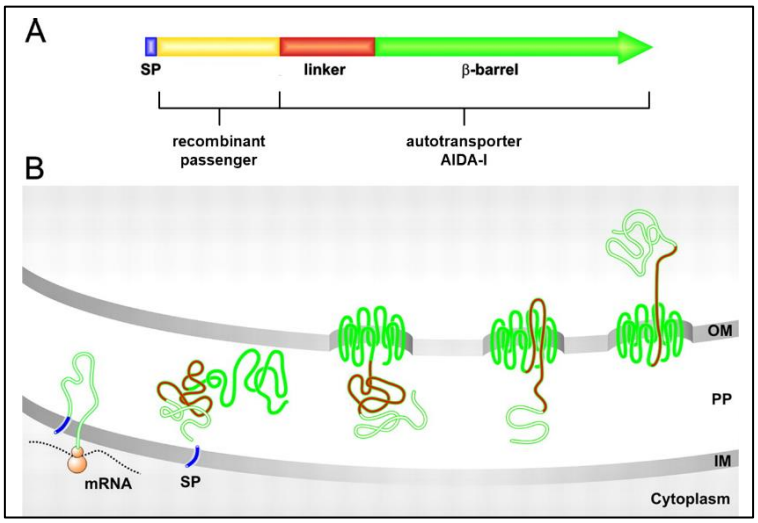

Figure 1.19: Autodisplay system for surface display in *E. coli*. (A) Schematic representation of the Autodisplay construct, consisting of a signal peptide (SP), a passenger domain and the translocator domain (linker and β-barrel) of the autotransporter AIDA-I. The passenger domain is replaced by the recombinant protein to be surface displayed. (B) Transport of the recombinant passenger. The signal peptide serves to facilitate transportation across the inner membrane (IM). In the periplasm (PP), the C-terminal translocator domain forms a porin-like structure (β-barrel) in the outer membrane (OM). Subsequently, the N-terminal passenger is translocated through the porin-like structure and is thereby surface displayed (figure modified from [248]).

Although it is not a class II enzyme, it might be worth noting that the self-sufficient *B. megaterium* CYP102A1 (119 kDa) was successfully surface displayed using three different systems (Figure 1.19 and Figure 1.20). In the first report, surface display was enabled by a truncated ice nucleation protein (INP) InaK of *Pseudomonas syringae* (Figure 1.20A) [250]. INPs are found in plant pathogenic bacteria and are composed of three distinct domains: a N-terminal unique domain containing the membrane anchor, a C-terminal unique domain and a central domain composed of repeats given by a 8-, 16- and 48-residue periodicity that acts as template for ice nucleation [251]. This central repeating domain is not required for transportation to and display on the outer membrane of the passenger (fused at the C-terminus

of INP) [252], [253] and was deleted in the truncated InaK construct. A second report on the surface display of CYP102A1 made use of the aforementioned Autodisplay system (Figure 1.19) [254]. Instead of the ctxB signal peptide, they used the gIII signal peptide, present in the pBAD gIII vector. In a recent publication, a SpyCatcher-SpyTag method was applied for the surface display of CYP102A1 (Figure 1.20B). The Spycatcher protein was fused to the Lpp-OmpA hybrid surface display anchor, i.e. the signal peptide and nine N-terminal residues of lipoprotein (Lpp) and the transmembrane domain of outer membrane protein A (OmpA). CYP102A1 was fused to the SpyTag peptide. Both proteins were produced separately in *E. coli*, after which the SpyCatcher protein and SpyTag peptide form a covalent isopeptide bond, thereby coupling CYP102A1 to a membrane anchor post-translationally, and enabling translocation to the *E. coli* surface [255].

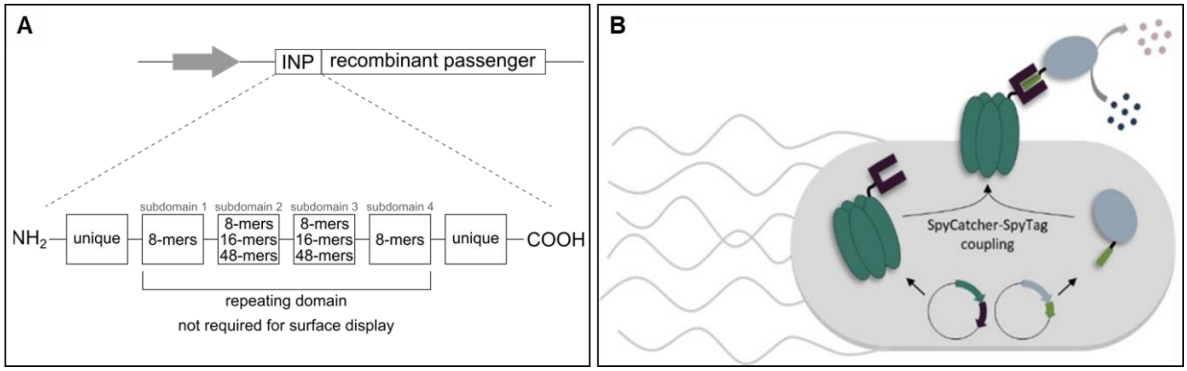

Figure 1.20: Two of three strategies for the surface display of CYP102A1 on the *E. coli* outer membrane. A: The recombinant passenger is fused to the C-terminus of the ice nucleation protein (INP) surface anchor. The central repeating domain of INP is not required for surface display and can be modulated. B: SpyCatcher-SpyTag method. CYP102A1 and the membrane anchor are independently produced and coupled post-translationally for surface display (figure taken from [255]).

For crystallization purposes, efforts were done to obtain a completely soluble enzyme and other engineering strategies were used. In order to decrease the hydrophobicity, the N-terminal membrane anchor is deleted, either partially or completely and in some cases a hydrophilic peptide is added [229], [230]. It has been reported that a proline-rich region is present following the N-terminal anchor, which is essential for the correct folding of the protein. The truncations should thus leave this region intact [256]. It should also be noted that deletion of the N-terminus might lead to loss of activity so care should be taken [229], [230]. Removal of the N-terminus alone might not be sufficient as it has been shown that the F-G loop region also interacts with the membrane and additional mutations might be required. For example, McDougle et al deleted the first 34 amino acids from CYP2J2 and substituted some hydrophobic amino acids in the F-G loop with hydrophilic amino acids. Upon substitutions in the F-G loop, more protein was retrieved in the cytosol compared to the membrane-truncated construct [257]. The very first crystal structure of a microsomal CYP2C5 is a good example of this as well (see 2.3). The sequence encoding the N-terminal anchor was removed, the second codon was altered to alanine and additional substitutions at the 5'-end were done in order to introduce a restriction site, resulting in the N-terminal sequence MAKKTSSKG [258]. Not only the N-terminal anchor was truncated, but amino acids within the F-helix and the F-G loop were substituted based on alignment with CYP2C3. This resulted in an enzyme, soluble and monomeric in high salt buffers [259]. Care should be taken when mutating the F-G loop as well as this is part of the

SRS regions. In the mutation studies with CYP2C5, they originally substituted residues 201 to 210 with these from CYP2C3 and observed a loss of activity. To reconstitute activity, some amino acids needed to be substituted back to the native residue. The short hydrophilic, positively charged peptide MAKKTSSKG was subsequently used for crystallization of other CYP2 family enzymes, whether or not including further truncations in the F-G region [260]–[262].

After initial production is established, other parameters could be optimized in order to obtain high yielding CYP. Frequently co-expression of the chaperones GroES/EL is performed, resulting in higher protein yields and 5-ALA is almost always added for increased production [229], [230]. Another factor is the choice of vector and host strain. The vectors pCWOri, containing a *tac* promotor, and the commercial pET vector with a the T7 promotor/*lac* operator are often used. In fact, the pCWOri vector appears to be the better choice for CYP production, although high yields have been reported with the pET vector as well [229], [230]. Several *E. coli* strains are commercially available and have been used for recombinant CYP production, JM109 often appearing as the strain of choice. However, as this strain does not carry a copy of phage T7 RNA polymerase, this cannot be used in combination with the pET vector [229], [230]. Of interest, high yields were reported as well when using the so-called Walker strain C41(DE3), a BL21(DE3) derivative for membrane protein overexpression [229]. Lastly, culture parameters could further alter the CYP yield, e.g. temperature, pH, rpm, addition of trace elements, point of induction, IPTG concentration, etc. [229], [230].

2.6.4 Providing the redox partner

2.6.4.1 Addition of a separate redox partner

CYPs from class II require one redox partner, i.e. CPR. In some cases an additional redox partner is involved, namely cytochrome *b5*. Eukaryotic expression hosts contain endogenous CPR and can sustain heterologous CYP activity. However, endogenous CPR might not be sufficient to sustain the overexpressed CYP and therefore, co-expression of the redox partner is often applied [229]. In yeast, this has been reported to lead to increased CYP stability [237]. In *E. coli*, no endogenous CPR is present and CPR co-expression is necessary for activity. However, it was shown that the endogenous flavodoxin and reductase of *E. coli* are able to reduce class II CYP enzymes [263]. Using whole cells for bioconversion, it has been observed that finetuning this co-expression is necessary, both in yeast and *E. coli*. In natural circumstances, CYP and CPR are not present in a 1:1 ratio and CPR sustains multiple CYPs, as well as other enzymes. In the production of artemisin in *S. cerevisiae* (see 2.1), it proved to be advantageous to express *cpr* from a weaker promotor [132]. Also the production of taxadien-5α-ol in *E. coli* (see 2.1) improved upon increasing the CYP:CPR ratio [129].

The first choice to select a redox partner for co-expression would be the native CPR redox partner. On the other hand, creating a single strain overexpressing one general CPR for the production of CYPs from different origins is often done, e.g. in screening systems. For example, a *S. cerevisiae* strain overexpressing its endogenous CPR, was used for a functional screening of a plethora of fungal CYPs [236]. Also in the creation of chimeric constructs, many examples

can be found where a reductase domain of a self-sufficient CYP was chosen in order to mimic the natural self-sufficient system (see further). However, care should be taken when selecting a redox partner. Using heterologous redox partners, it was seen that these could majorly influence the activity. For example, a selection of human CYPs were co-expressed in *S. pombe* with human CPR, CPR from the plant *Ammi majus* and the endogenous CPR from *S. pombe*. While human CPR proved to be the better partner for the CYP3 family, this did not necessarily hold true for others. For two human CYPs, they saw that not human CPR but one of the other CPRs resulted in the highest activity [264]. Fungal CYP53A19 of *F. oxysporum* performed best with its endogenous redox partner, converting the substrates benzoic acid, 3-methoxybenzoic acid and 3-hydroxybenzoic acid. Both benzoic acid and 3-methoxybenzoic acid, but not 3-hydroxybenzoic acid, were converted when using *S. cerevisiae* CPR and only benzoic acid was hydroxylated in combination with *C. albicans* CPR [265]. In another example, Chang et al reported that fusion of the fungal CYP57B3 from *Aspergillus oryzae* to the reductase domain of CYP102A1 mainly resulted in 6-hydroxydaidzein. When using the native redox partner, production of 8-hydroxydaidzein was reported and with the CPR from *S. cerevisiae*, a 3'-hydroxyl derivative was produced [266]. A remarkable example, be it not from a class II-type CYP, is given by Zhang et al, who observed that separately adding the reductase domain of the class VII CYP116B2 to the bacterial CYP MycG resulted in a completely different reactivity [267].

In CYP characterization studies, CPR enzymes can be added separately, either isolated from their natural host or recombinantly produced and purified. Similar to CYP, CPR is a membrane protein, integrated in the microsomes with an N-terminal anchor. However, the hydrophobicity of the enzyme is lower than CYP and deleting this anchor appears sufficient to obtain a soluble enzyme. There is still some controversy about the function of this anchor, as discussed before.

2.6.4.2 Chimeric constructs

Inspired by the high catalytic activity of the self-sufficient CYP102A1 from *B. megaterium*, together with the high coupling efficiency, many chimeric enzymes have been constructed, fusing a microsomal CYP with a microsomal CPR (Figure 1.21A). In this way, the heme, FMN and FAD domain are found in the same polypeptide chain, eliminating the need for two separate enzyme partners. Only NADPH and the substrate need to be added in this case. The very first example of such a construct was produced in 1987 by Murakami et al [268], where rat CYP1A1 was fused to rat CPR using a short (three amino acids) linker. The CPR was N-terminally truncated in order to remove the membrane anchor, 56 amino acids were removed in accordance to the trypsin digested construct for crystallization (see 2.5). CYP was not N-terminally modified and expression was performed in *S. cerevisiae*. Both the isolated microsomes and the purified chimer showed to be active. The same group extended its research by making fusion constructs between bovine CYP17A1 and *S. cerevisiae* CPR, where they tested alternative CPR truncations with or without a three amino acid linker. The CPR where only part of the N-terminus, cleaved of by papain digest, was removed, proved to be the most active construct [269]. The group of Estabrook used the same approach to construct chimers between microsomal CYPs and truncated microsomal CPRs, using a Ser-Thr linker [270]. In this case,

also the N-terminus of CYP was modified in order to enable *E. coli* expression. This group and others continued to make other constructs using similar approaches, which have been reviewed comprehensively in [271]. Varying success has been achieved with this approach. Some chimers indeed showed increased activity upon purification together with good coupling efficiencies. However, not all chimers were equally successful. Lipid and detergent addition was required for activity and cytochrome *b5* was necessary for increased coupling efficiency in several cases. Due to the necessity of cytochrome *b5* in some instances, three-protein-fusions have been created as well [272]. Next to the mammalian chimers, also other fusion proteins using class II CYPs have been constructed, e.g. between *Thlaspi arvensae* CYP71B1 and *Catharanthus roseus* CPR and between *C. roseus* CYP71D12 and its CPR [271]. The recent example of CYP52M1 from *S. bombicola* fused to CPR from *A. thaliana* in an engineered *S. cerevisiae* strain for OHFA production further illustrates how creating a chimeric construct can improve catalytic properties. Of note, the heterologous redox partner CPR from *A. thaliana* resulted in higher conversion yields over the homologous redox partner [58].

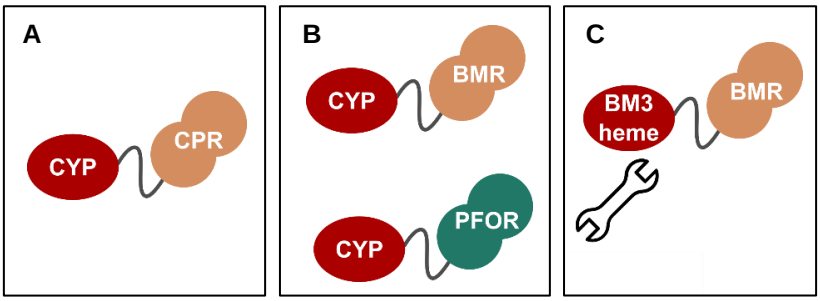

Figure 1.21: Different approaches for the construction of a self-sufficient CYP. A: fusion of the class II CYP to a class II redox partner CPR. B: fusion of the reductase domain of a natural self-sufficient CYP to a class II CYP. C: Protein engineering of a natural self-sufficient CYP.

Another approach was proposed by Gilardi et al, where a so-called molecular lego method was applied [273]. Instead of fusing microsomal CYPs to a microsomal redox partner, the reductase domain of CYP102A1 was introduced C-terminally of a truncated CYP (Figure 1.21B). This technique showed to be successful in creating a self-sufficient system for several mammalian microsomal CYPs and no additional proteins were needed. On top, solubility was increased and the enzyme was active in the absence of detergents and lipids [271]. Further studies showed the importance of the linker regarding coupling efficiency, catalytic activity and substrate binding [274], [275]. It was seen that using a longer five glycine-residue linker resulted in improved properties. The reductase domain of CYP116B2 of class VII has emerged as a possible fusion partner as well (Figure 1.21B). This was already exemplified in the ω-hydroxylation of FAs by the bacterial CYP153 and showed promise for bacterial CYPs [276]. This reductase domain showed to support the activity of eukaryotic CYPs upon fusion as well [277].

On the one hand, artificial chimers can be constructed. On the other hand, natural chimers could be used in biotechnological applications, as self-sufficient CYPs have been identified throughout the years. These have been reviewed more recently in [137]. Current research thus also focusses on engineering these natural chimers (Figure 1.21C), either by rational design or directed evolution in order to obtain foreseen activity and selectivity. This has been exemplified by engineering CYP102A1 to enable selective ω-oxidation of FAs (see 1.2.2).

2.6.4.3 Other electron delivery systems

In both approaches described above, obtaining high coupling efficiencies remains challenging. Additionally, NADPH is required in stoichiometric amounts, unless cofactor regeneration is in place. Cofactor regeneration is established metabolically in whole-cell biocatalysts and some enzymes have been put forward for *in vitro* cofactor regeneration (see further). Other ways of circumventing the need for NADPH and/or a redox partner in CYP catalysis have been investigated.

A first example has already been mentioned before (see 2.4), i.e. making use of the peroxide shunt pathway. In this case, instead of using oxygen together with electron delivery from NAD(P)H, hydrogen peroxide is used by CYP. The high spin species formed upon substrate binding is then immediately converted to the ferric hydroperoxo intermediate Compound 0. Also other so-called oxygen surrogates were applied in this context, such as iodosylbenzene, sodium chlorite, cumene hydroperoxide and m-chloroperbenzoic acid [176], [278]. Remember, m-chloroperbenzoic acid was used in order to spectroscopically characterize Compound I (see 2.4). Several examples are found in which mammalian CYPs could use one or more of these oxygen surrogates [279]–[282]. However, stability issues and heme inactivation arises when using hydrogen peroxide. Additionally, most CYPs need to be mutated in order to use oxygen surrogates in an efficient way. The group of Frances Arnold, for example, created a CYP102A1 mutant by directed evolution, able to use hydrogen peroxide for the hydroxylation of *p*-nitrophenoxydodecanoic acid, FAs and styrene [283]. Notably, some CYPs evolved to be actual peroxygenases and use hydrogen peroxide as the oxidant [278].

An electrochemical approach where CYPs are immobilized on electrodes, serves as a potential alternative. Electrodes can deliver electrons directly to the heme. In order to do this efficiently, the electrode surface is best engineered. Different electrode materials have been used, such as graphite, glassy carbon, gold and other metals. Many surface modifications have been applied for enzyme immobilization. The surface has been coated with surfactants, polymers, gold nanoparticles and self-assembled monolayers. Stable vesicular phospholipid dispersions have been used in order to create a lipid film and microsomes have been immobilized on the electrodes. The immobilization strategy, involving also enzyme orientation, the use of spacers, etc, greatly affects the stability and selectivity of the enzyme. Also protein engineering has been proposed and the chimeric constructs, as described above, have been applied [284], [285]. This electrochemical approach is of particular interest in the context of biosensors [286]. However, heme reduction by the electrode was stated to be insufficient in context of industrial biocatalysis and issues concerning CYP stability upon electrode fixation need to be solved [122].

More recently, light-driven electron delivery has been put forward as a promising strategy. The idea is that a photosensitizer transfers light energy to the chemical reaction for electron supply. Different photosensitizers have been under investigation in CYP catalysis; quantum dots, deazaflavins, isoalloxazine flavins riboflavin, FAD and FMN, eosin Y, Ru(II)-diimine complexes and photosystem I (remember redirecting the metabolic pathway to the chloroplast of plants and the use of cyanobacteria). Not all photosensitizers deliver the electrons in the same

way. Firstly, some photosensitizers transfer electrons to oxygen, resulting in peroxide which than continues through the peroxide shunt pathway. Secondly, photosensitizers deliver electrons to the heme via a redox partner. Thirdly, no redox partner is involved an the photosensitizer directly reduces the heme group. These light-driven reactions have recently been reviewed in [287]. Such systems still require a lot of optimization regarding the stability and activity of the CYP enzymes and in case of the direct heme reduction, coupling efficiency needs to be optimized as well.

2.6.5 Cofactor regeneration

To be able to use CYPs on an industrial scale, cofactor regeneration is required. In CYP biocatalysis, several alternative electron delivery systems have been investigated in order to circumvent the need for NAD(P)H, as discussed above. An additional advantage of many of these systems is the fact that often no redox partner is required. However, none of these alternative electron delivery systems has yet found its way to industrial application and for cofactor regeneration on an industrial scale, enzymatic regeneration is still the most common approach [120], [122].

Cofactor regeneration is established metabolically in whole-cell biocatalysts and this is one of the reasons why whole-cell bioconversion is mostly applied in case of CYP-mediated catalysis. However, intracellular NADPH levels can become rate limiting when CYPs are highly active. To overcome this limitation, cofactor regeneration systems are co-expressed. In several cases, glucose dehydrogenase (GDH) has been co-expressed with engineered CYP102A1 enzymes, for example in [288], [289].

For *in vitro* applications, cofactor regeneration is a must and several dehydrogenases have been used for this purpose. Glucose-6-phosphate dehydrogenase (G6PDH) has a high specific activity and is highly selective for the phosphorylated cofactor $NADP^+$. G6PDH is used frequently for cofactor regeneration on a laboratory scale [290], [291]. However, the substrate glucose-6-phosphate (G6P) is expensive and therefore G6PDH is not of use on an industrial scale. Alcohol dehydrogenase (ADH) was proposed as a useful NADPH regenerating enzyme as the alcohol substrate can serve as a solvent for the hydrophobic CYP substrate [290]. For example, Kubo et al used the ADH from *Thermoanaerobium brockii* for NADPH regeneration in the epoxidation of terminal alkenes using a mutant CYP102A1 [292]. Alternatively, formate dehydrogenase (FDH) is of specific interest for cofactor regeneration on an industrial scale, because of the many advantages. The substrate formate is cheap, readily available and innocuous to most enzymes. The reaction is irreversible and the formed product CO_2 is easily removed, thereby not interfering with downstream purification processes. One of the disadvantages is that it has a specificity for NAD^+ as cofactor. However, protein engineering of the FDH from *Pseudomonas* sp. 101 has resulted in a $NADP^+$-specific FDH, able to recycle NADPH in the conversion of β-ionone, octane and naphthalene by CYP102A1 mutants [293]. More recently, also phosphite dehydrogenase (PTDH) has gained attention as a cost efficient regeneration system. The substrate phosphite is cheap and the reaction is thermodynamically almost irreversible. The enzyme can accept $NADP^+$, although mutations are required to increase

$NADP^+$ over NAD^+ specificity. Beyer et al showed how fusion of such an engineered PTDH to CYP102A1 resulted in a true self-sufficient enzyme, including an efficient cofactor regeneration system in the same polypeptide chain [294].

NADH is less expensive and more stable than NADPH [295]. Additionally, NADH regeneration systems are more readily available. This has prompted researchers to investigate whether the cofactor specificity of the CYP redox partner could be switched. The self-sufficient CYP102A1 proved of value in these investigations as well and Maurer et al successfully engineered the reductase domain of CYP102A1 in order to accept NADH instead of NADPH [296].

2.6.6 Immobilization for *in vitro* application

Industrial application of CYPs of class II is now limited to whole-cell bioconversion approaches for several reasons such as enzyme instability and their requirement for redox partner(s) and cofactor. Using a whole-cell biocatalyst comes with some disadvantages as well which already emerged during the discussion of DCA production by fermentation. Interfering enzymes and pathways might be present, resulting in unwanted side reactions and/or byproducts. The substrate and/or product can be toxic for the cells. Moreover, limitations regarding substrate uptake and/or product secretion occur. Many strategies have been investigated in order to overcome these disadvantages. Strain engineering can be performed to knock out unwanted side reactions, transporters can be introduced, cells can be permeabilized for increased substrate uptake, a two-phase system can be set up, etc. [120]. Alternatively, an *in vitro* cell-free approach delivers a solution to these disadvantages. This offers an additional advantage for implementation in chemical industries, as whole-cell biotransformation requires specialized equipment, not readily available in chemical industry. In case of *in vitro* biocatalysis, no such equipment is necessary. Unfortunately, CYPs are known to be unstable in a soluble form. Immobilization of an enzyme for increased stability is a well-known strategy and also for CYP enzymes, immobilization has been explored. Several approaches for immobilization have been described in literature, both as a research tool and for biocatalytic application. Immobilization on electrodes, described above, is already one example of this.

Eukaryotic CYPs are naturally immobilized in the ER membrane. The isolation of microsomes is thus an option in obtaining CYPs for *in vitro* application in an immobilized form. Indeed, for activity assays this is often used. However, microsome preparation is laborious and interference of other CYPs limits their use. The technically demanding purification protocol and instability also limits their use on an industrial scale [114].

Early attempts to immobilize class II CYPs were reported to make use of calcium alginate and BrCN-activated sepharose 4B for the entrapment of *S. cerevisiae* CYP [297]. Maurer et al showed that sol-gel matrix could be used for the immobilization of CYP102A1 together with FDH for cofactor regeneration. Improved stability was observed, tested over a period of 30 days, and catalyst recycling was achieved [293]. Co-immobilization of CYP102A1 and the cofactor regeneration enzyme (i.e. GDH) has been reported by Valikhani et al as well. They

accomplished this by fusing a binding module Z_{basic2} to both CYP102A1 and GDH for immobilization on an anionic sulfopropyl-activated carrier (ReliSorb SP) [298]. In another study, plant CYP71B1 and its CPR were fused and incorporated in colloidal liquid aphrons (oil-in-water macro-emulsion) and showed a 10-fold improved activity [299].

In membrane protein research, detergent solubilization and the incorporation in liposomes has long been established. CYP and CPR, being integral membrane proteins, have been incorporated in phospholipid vesicles for functional assays [300], for example for the investigation of membrane composition on enzyme function [301]. More recently, so-called Nanodiscs have emerged as an alternative to these detergent micelles and liposomes. These Nanodiscs are lipid bilayers, discoidal in shape, encircled by an amphiphatic belt (Figure 1.22). This belt is formed by helical scaffold proteins, derived from apolipoprotein A-I. These membrane scaffold proteins can be adjusted as a way of tuning the particle size. Nanodiscs are monodisperse, reproducible and relatively robust and stable [302]. They have found their way in CYP and CPR research, which has been reviewed in [303] and more recently in [304].

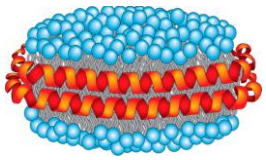

Figure 1.22: Nanodisc, a dicoidally shaped lipid bilayer, encircled by an amphiphatic belt (figure taken from [304]).

Many microorganisms have been reported to produce polyhydroxyalkanoates (PHA), of which polyhydroxybutyrate (PHB) is the most abundant polyester. These compounds are produced for storage of carbon and energy. PHAs are found intracellularly as granules, surrounded by a layer of proteins and phospholipids [305]. Lee et al found that CYP102A1, fused to the PHB-associated protein phasin, was incorporated at the granule membrane surface in *E. coli* (Figure 1.23). This PHB-immobilized CYP102A1 showed increased yield and stability over the free enzyme [306]. The group of Hannemann used *Bacillus megaterium* as a recombinant host for the expression of the mitochondrial CYP11A1 and showed the enzyme to be localized at the surface of the PHB granule, without a phasin-tag. The enzyme was active and showed a significantly increased stability, leading to the first mammalian CYP to be lyophilized [307]. Although this does not concern class II CYPs, it might serve as an immobilization carrier for microsomal CYPs as well.

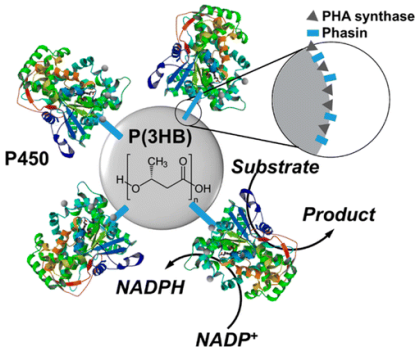

Figure 1.23: Polyhydroxybutyrate, here indicated as P(3HB), granule, isolated from *E. coli*. CYP102A1 is integrated in the membrane surface using the PHB-associated protein phasin as the membrane anchor (figure taken from [306]).

Chapter 2
Aim of the project

This PhD was started as part of the project Enzymes for Added Sustainability and Efficiency (EnzymASE). The project was funded by The Flemish Agency for Innovation and Entrepreneurship (VLAIO) and supported by Catalisti. EnzymASE is a cooperation between different industrial partners and knowledge centres, including Ghent University. Its aim is to develop enzymes and corresponding biocatalytic processes for a variety of applications; novel enzyme tools, feed additives, coatings and adhesives. One of the industrial partners was interested in saturated long-chain α,ω-DCAs for the use in repellent materials. Developing a sustainable production process for these compounds was known to be challenging and was therefore presented as a PhD project. Long-chain α,ω-DCA production is not only of interest in this specific application. In a broader context, these compounds are of value as a monomer in polyamides and polyesters and as precursors for pharmaceuticals, perfumes, lubricants, adhesives, plasticizers, powder coatings and corrosion inhibitors. Using long-chain α,ω-DCAs gives attractive properties to the polymers compared to currently used short-to-medium-chain DCAs, such as decreased water absorbance, which affects stability and mechanical performance, and more polyethylene-like properties are obtained, i.e. higher melting temperatures and a higher tendency to crystallize. Moreover, these long-chain DCAs can lead to entirely new polymers.

Two different processes are in place in industry for the production of long-chain α,ω-DCAs. Chemical conversions are industrialized, of which only the metathesis of monounsaturated FAs delivers long-chain α,ω-DCAs. This chemical conversion process is accompanied with a number of disadvantages. Byproduct formation occurs by cross-metathesis reactions when the substrate is a mixture of monounsaturated FAs, and by isomerization. This leads to an extensive purification process, elevating the costs. Only half of the FA substrate is included in the final product and expensive rare catalysts are used for the chemical conversion, which cannot be recovered and contaminate the final product. Lastly, saturated DCAs are of specific interest in this project so this chemical conversion would require an extra hydrogenation step. Many of these disadvantages are overcome in the second production process in industry, that is the fermentation process using *Candida viswanathii*. Unfortunately, this organism is classified as pathogenic in Europe (risk group class two) and the necessary safety precautions that need to be taken, leads to increased prices. Additionally, both approaches are protected by patents [308], [309]. A sustainable production process was pursued, which is not only lower in costs, but is also more flexible and yields products with higher purity and consistency. An *in vitro* approach was of specific interest. This circumvented issues regarding substrate uptake limitations of long-chain FAs and unwanted side reactions in the cellular environment (e.g. degradation of FAs and DCAs via β-oxidation and the inclusion of FAs in lipid bodies), encompassed in fermentative approaches. Furthermore, often higher productivities are obtained in an *in vitro* process and products do not need to be recovered from the complex fermentation broth, enabling higher purities. Additionally, integration of a biocatalytic process in established chemical companies is more straightforward, as they have little know-how on fermentative production processes and huge investments regarding installations would be required for fermentation.

Two strategies were considered in this dissertation for the development of an *in vitro* cell-free production process (Figure 2.1A and B, respectively). In a first strategy, the recombinant production of a soluble variant of CYP52A13 from the industrial DCA producer *C. viswanathii* was investigated. A soluble enzyme allows for a facilitated implementation in the enzyme reactor and provides a flexible system. In a second strategy, *in vivo* immobilization of the same biocatalyst on the surface of *Escherichia coli*-derived outer membrane vesicles (OMVs) was explored. It was hypothesized that immobilization on these OMVs would enhance enzyme stability. Moreover, OMVs might be recovered after biocatalysis for recycling, important for reducing the costs of the overall production process. In both strategies, the production of a self-sufficient CYP was pursued. Creating a chimeric construct where both the CYP and the redox partner CPR are encoded in the same polypeptide chain, eliminates the necessity of producing two individual enzymes. Additionally, it has been reported that fusing CYP and CPR can significantly enhance the coupling efficiency. Furthermore, cofactor regeneration using FDH was addressed. As NADPH is an expensive cofactor, large-scale application of this system would not be feasible if the NADPH has to be added in stoichiometric amounts. FDH was chosen as it is irreversible, requires an inexpensive co-substrate and the formed CO_2 is inert and can be easily removed.

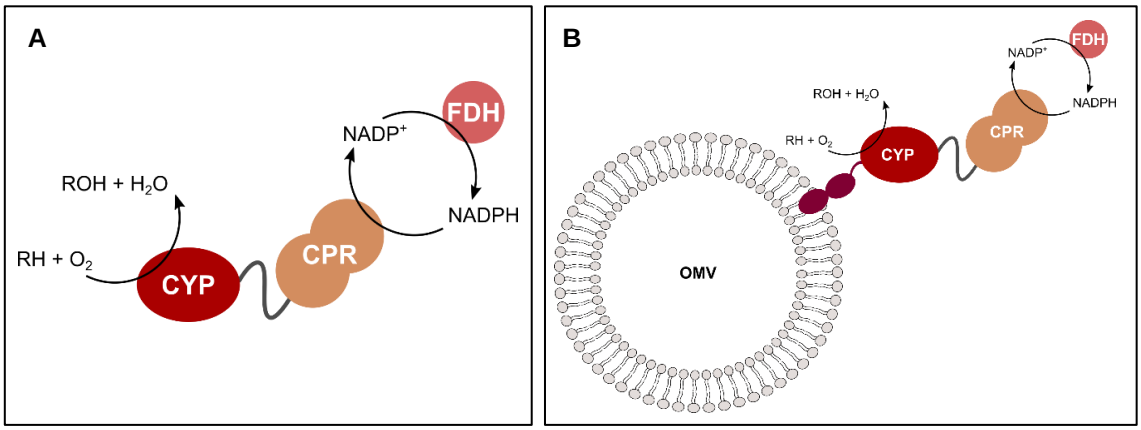

Figure 2.1: Proposed *in vitro* systems. A: soluble self-sufficient CYP, produced in *P. pastoris*. B: self-sufficient CYP, produced in *E. coli* and displayed on *E. coli*-derived outer membrane vesicles (OMV). Cofactor regeneration is pursued by recombinantly producing the FDH from *Burkholderia stabilis*.

The first experimental chapter (chapter 3) describes the recombinant production of a soluble self-sufficient CYP using the heterologous host *Pichia pastoris*. On the one hand, both CYP52A13 and CPR-a were produced separately in a soluble and secreted form, after which a chimeric construct was made. On the other hand, the nontruncated CYP52A13 was produced, intended for optimization of the chimeric construct in the native ER environment, before continuing to further solubilization efforts. The second experimental chapter (chapter 4) reports the search for an *in vivo* immobilization strategy using OMVs. The natural self-sufficient CYP102A1 from *Bacillus megaterium* was used as a model self-sufficient CYP in order to deliver a proof-of-concept. In a next step, the self-sufficient CYP52A13 was first surface displayed and subsequently displayed on OMVs. The third chapter (chapter 5) then focusses on the recombinant production of FDH for cofactor regeneration. The $NADP^+$-dependent FDH from *Burkholderia stabilis* was selected for recombinant production in two different hosts, i.e. *P. pastoris* and *E. coli*.

Chapter 3
A soluble self-sufficient CYP for *in vitro* production of long-chain α,ω-DCA

Delphine Devriese[1], Stijn De Waele[1], Jordy Bauwelinck[2], Pieter Surmont[3], Sören Planckaert[1], Iris Cornet[2], Frederic Lynen[3], Bart Devreese[1]

[1]Laboratory of Microbiology – Protein Research Unit, Ghent University
[2]Biochemical Wastewater Valorization and Engineering (BioWaVE), University of Antwerp
[3]Separation Science Group, Ghent University

Authors contribution: Delphine Devriese wrote the whole chapter and performed all experiments. Stijn De Waele designed the truncated CYP and CPR constructs. Jordy Bauwelinck performed GC-MS of the methylated samples. Pieter Surmont performed GC-MS of the silylated samples. Sören Planckaert did the MRM analysis. Iris Cornet and Frederic Lynen are the supervisors of Jordy Bauwelinck and Pieter Surmont, respectively. Bart Devreese is the supervisor of the author.

Abstract

Long-chain α,ω-dicarboxylic acids (DCAs) are valuable chemical compounds with many applications, e.g. as a monomer in polyamides and polyesters. Unfortunately, only a limited set of long-chain DCAs is commercially available, produced either by chemical conversion of fatty acids (FAs) through metathesis or by fermentation using *Candida viswanathii*. Moreover, both processes come with several disadvantages so an alternative long-chain DCA production process was pursued in this dissertation. Cytochrome P450 monooxygenases (CYPs) are versatile enzymes with a potential in a variety of synthetic biology applications. The ability to introduce one oxygen atom from molecular oxygen in a non-activated C-H bond while the other oxygen atom is reduced to water, makes them attractive biocatalysts. In addition, they are able to perform these reactions in mild reaction conditions and often exhibit high regio-, chemo- and/or stereoselectivity. A number of CYPs are reported to regioselectively hydroxylate the terminal chain position of FAs and overoxidation to the corresponding DCA has been reported. From these, CYP52A13 from *C. viswanathii* was selected and it was investigated if it was feasible to create an enzyme reactor for long-chain α,ω-DCA production. To this end, both CYP52A13 and the redox partner CPR-a were truncated in order to produce them either individually in a separate host or as one chimer in a soluble and secreted form. In parallel, a nontruncated form of CYP52A13 was produced in *Pichia pastoris*, intended for optimization of the biocatalyst. Unfortunately, low yields were obtained and activity towards the substrate oleic acid could not be confirmed. However, adding an N-terminal tag led to increased CYP yields and showed to support secretion to the medium, unveiling a potential approach for the development of CYP based reactor systems.

1 Introduction

Long-chain α,ω-dicarboxylic acids (DCAs) are valuable chemical compounds with applications as a monomer in polyamides and polyesters and as precursors for pharmaceuticals, perfumes, lubricants, adhesives, plasticizers, powder coatings and corrosion inhibitors [1]–[3]. Today, only a limited set of long-chain DCAs is commercially available [1]–[3]. Two production processes are in place in industry, i.e. a chemical conversion of fatty acids (FAs) by metathesis [13] and a fermentation process using (engineered strains of) the opportunistic pathogen *Candida viswanathii* [26]. Both processes come with a number of disadvantages and the market price is elevated. As the use of long-chain DCAs is increasing, there is a demand for alternative, cost-effective production processes [1]–[3].

Cytochrome P450 monooxygenases (CYPs) are versatile enzymes with a great potential in a variety of synthetic biology applications [120]–[124]. CYPs contain a b-type heme group, linked to the enzyme via a strong bond between the heme iron and a conserved cysteine residue. The structure around this heme is highly conserved. This is in line with their common catalytic mechanism, where they introduce one oxygen atom from molecular oxygen in a non-activated C-H bond while the other oxygen atom is reduced to water [146], [169]. This reaction is difficult to achieve by conventional chemical routes, which often require harsh conditions using rare metal catalysts. In contrast, CYPs are able to perform these reactions in mild reaction conditions and often exhibit high regio-, chemo- and/or stereoselectivity, making them attractive biocatalysts [120]–[124]. Even though all CYPs share the same unique P450 fold, consisting of a helix-rich α-region and a helix-poor β-region, some sites show high structural flexibility, enabling a high variability of substrates within this superfamily [146], [169].

A number of CYPs have been reported to selectively hydroxylate alkane and/or FA substrates at the ω-position, i.e. the ER-bound fungal CYP52 family, mammalian CYP4 family and plant CYP76, CYP78, CYP86, CYP94 and CYP96 families. Also the soluble bacterial CYP153 family has been shown to perform this ω-hydroxylation [71]. Furthermore, ω-oxidizing CYPs (ER bound), other than the CYP52 family, were identified in filamentous fungi, such as CYP539A7 and CYP655C2 from *Fusarium oxysporum* [114]. A number of these CYPs have been investigated in the context of industrial DCA production. Fermentation processes have been the main focus in this field thus far, with the CYPs either expressed in their natural (engineered) host or in a heterologous host [2], [3]. This is primarily due to the fact that they require a redox partner and the cofactor NAD(P)H which can be delivered by the host. Moreover, reactive oxygen species (ROS) are often formed during CYP catalysis and the cellular environment contains enzymes able to decompose these species [120], [122]. However, fermentation comes with some disadvantages. Strain engineering is necessary in all cases in order to delete the β oxidation pathway, which is not always that straightforward. Substrate uptake as well as secretion of the end product can be a major limiting factor, especially when long-chain DCAs are of interest. Additionally, toxicity of the substrate and/or product also can occur, e.g. in the case of medium-chain FAs [2], [3].

Aforementioned disadvantages could be overcome by applying an *in vitro* approach where a purified enzyme could be introduced in a reactor system. Based on the fact that yields obtained by fermentation using the yeast *C. viswanathii* are unmatched until today, enzymes from this species were selected to develop an enzyme-based approach to synthesize long-chain α,ω-DCAs, which are the most challenging to produce. Craft et al showed that the genes encoding CYP52A13 and CYP52A17 were induced upon oleic acid addition in *C. viswanathii* [23]. Eschenfeldt et al then showed that both CYP enzymes were able to oxidize the longer oleic acid substrate and overoxidation to the corresponding DCA was observed. CYP52A17 has a broader substrate specificity with a preference towards shorter chain saturated FAs, whereas CYP52A13 is more selective towards the longer chains [24]. Therefore, CYP52A13 from strain ATCC 20336 (formerly classified as *C. tropicalis*) was chosen for recombinant production in the heterologous host *Pichia pastoris*. CYP52A13 requires the redox partner NADPH-dependent cytochrome P450 reductase (CPR), and is thereby classified as a class II-type CYP, according to Hanneman et al [134]. The genome of *C. viswanathii* (ATCC 20336) contains two allelic variants coding for CPR, i.e. *cpr-a* and *cpr-b*. He and Chen determined the nucleotide sequences of two *cpr* variants in *C. tropicalis* strain 1230 and the allelic variants show high similarity (96.6 % identity between the ORFs). The two respective protein products, CPR-a and CPR-b, display 98.7 % amino acid sequence identity and their activity proved to be similar. They inferred that these nucleotide sequences were the same as the sequence of the allelic genes from *C. viswanathii* ATCC 20336 [310]. CPR-a was chosen for heterologous expression, as this variant was used by Eschenfeldt et al for supporting CYP52A13 and CYP52A17 activity [24]. Both CYP52A13 and CPR-a are bound to the ER membrane with an N-terminal anchor. Although *P. pastoris* contains the required organelle being an eukaryotic microorganism, it was of interest to obtain the enzymes in a soluble and secreted form, enabling a facilitated implementation in a reactor system.

The yeast *P. pastoris* was chosen for recombinant production and secretion of the enzymes as this offers several advantages. The yeast can be easily cultivated on inexpensive media and molecular techniques are well established. *P. pastoris* fermentation processes on an industrial scale are already developed long-time, first being cultivated as single cell protein in the 1970s by Philips Petroleum and since the 1980s used as a recombinant protein production host. Furthermore, the yeast possesses the GRAS status (Generally Regarded As Safe). It is a Crabtree negative microorganism, allowing for growth to high cell densities and heterologous expression is tightly controlled using the AOX1 promotor, which is also highly inducible. Lastly, the *P. pastoris* has a high secretory capacity. Little endogenous protein is secreted so secretion to the medium already serves as a first purification step, greatly facilitating subsequent downstream processing [311]–[314].

On the one hand, both CYP52A13 and its redox partner CPR-a were truncated in order to produce both enzymes either individually in a separate host or as one chimer in a soluble and secreted form. On the other hand, a nontruncated form of CYP52A13 was produced in *P. pastoris*, intended for optimization of the chimer and of the enzymatic assays. Unfortunately, no active biocatalyst was obtained. A flowchart displaying the different issues is depicted in Figure 3.1.

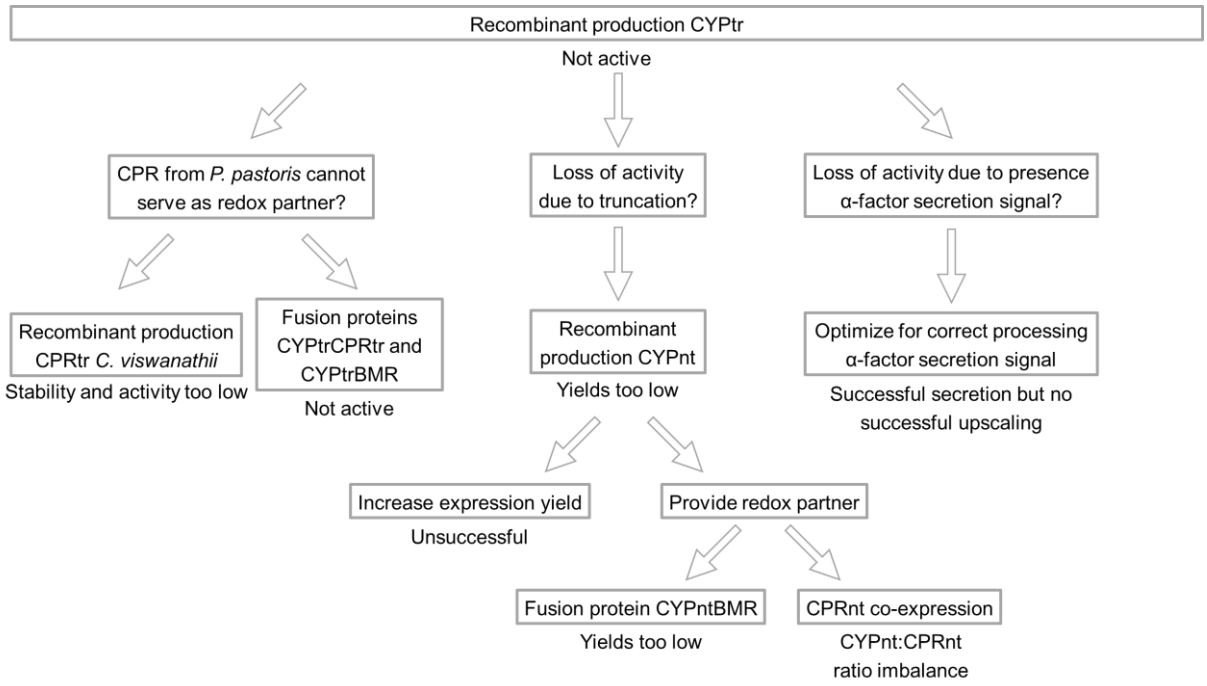

Figure 3.1: Flowchart of Chapter 3, illustrating the rationale of the different experiments.

2 Material and methods

2.1 Materials

All chemicals were purchased from Sigma-Aldrich, unless specified otherwise.

2.2 Strains and media

For cloning and plasmid amplification in *Escherichia coli*, either XL1-blue (Agilent) or DH5α (New England Biolabs (NEB)) was used. *E. coli* cells were grown in low-salt Luria-Bertani (LSLB) medium, including 50 µg/ml Zeocin™ (InvivoGen). LSLB consists of 0.5 % (w/v) NaCl (Merck), 0.5 % (w/v) yeast extract (Lab M) and 1 % (w/v) tryptone (Lab M). For agar plates, 1.5 % (w/v) agar (Lab M) was included. In case of CPRnt_pAO815, LB medium was used, containing 1 % (w/v) NaCl instead of 0.5 % (w/v) and including 100 µg/ml carbenicillin (Cb) (Gold Biotechnology).

Recombinant protein production was performed in the *P. pastoris* (also known as *Komagataella phaffii*) strain NRRL-Y-11430 (kindly provided by Prof. Nico Callewaert (VIB, Ghent University)). Alternatively, a *HAC1* co-expressing strain was used, where an expression construct for the spliced form of *HAC1* was transformed in *P. pastoris*, as described elsewhere [315]. This strain was kindly provided by Prof. Nico Callewaert (VIB, Ghent University). For the agar plates and precultures, yeast extract peptone dextrose (YPD) medium, including 100 µg/ml Zeocin™ (InvivoGen) was used. In case of *HAC1* co-expression, 300 µg/ml Hygromycin B Gold (InvivoGen) was included as well. YPD consists of 1 % (w/v) yeast extract (Lab M), 2 % (w/v) peptone (BD), 2 % (w/v) glucose, including or excluding 1.5 % (w/v) agar (Lab M), respectively. The growth phase was performed in buffered glycerol-complex medium (BMGY), consisting of 1 % (w/v) yeast extract (Lab M), 2 % (w/v) peptone (BD), 100 mM potassium phosphate buffer, pH 6, 1.34 % (w/v) yeast nitrogen base (YNB, Formedium) and 1 % glycerol (v/v). The expression phase was performed in buffered methanol-complex medium (BMMY). The composition is the same as for BMGY, except that 1 % methanol (v/v) (VWR) is added instead of 1 % glycerol.

2.3 DNA constructs

For both CPR-a and CYP52A13 from *C. viswanathii*, a truncated form, further referred to as CPRtr and CYPtr, respectively, was designed as described below. Codon optimization, gene synthesis and cloning was outsourced and performed by GenScript. The constructs were cloned into the vector pPICZαB in frame with the α-factor secretion signal in between the restriction sites PstI and SalI, present in the multiple cloning site of pPICZαB. The C-terminal Histag in the vector was excluded by keeping the stop codon. The constructs are denoted as CPRtr_ pPICZαB and CYPtr_pPICZαB, respectively. The expression vectors and codon optimized sequences are found in Addendum.

CPR-a was truncated in a similar fashion as described by Venkateswarlu et al for the truncation of the homologous CPR of *S. cerevisiae* (yCPR) [316]. Both protein sequences were aligned and the same region was deleted, up until the Phe36 (Figure 3.2). The artificial sequence used by Lamb et al [193], i.e. a Histag of six histidines, preceded by four and followed by 15 amino acids, which contained a thrombin cleavage site (MGSSHHHHHHSSGLVPRGSHMLDIM), was added N-terminally and the residues Met and Ser were added to substitute Phe. A last adaption was performed, as the codon CUG encodes Ser instead of Leu in *C. viswanathii* [24]. Therefore, this codon was modified for heterologous expression in *P. pastoris*.

Figure 3.2: Alignment of *C. viswanathii* CPR-a (tr|Q66T17|Q66T17_CANTR) to *S. cerevisiae* yCPR (sp|P16603|NCPR_YEAST). Only the N-terminus is shown. The arrows indicate the site of truncation.

CYP52A13 was N-terminally truncated according to the secondary structure prediction obtained using Phyre2 [317]. 99 % of the residues were modelled at more than 90 % confidence. However, the first seven residues were modelled *ab initio*, leaving uncertainties about the reliability of the predictions for this part of the protein. An N-terminal helix is formed by the first 40 amino acids and includes the transmembrane helix (Figure 3.3). It was decided to delete this helix, up until Tyr38. The N-terminal tag described for CPR was included at the N-terminus. Additionally, some unconserved phenylalanine residues near the N-terminus were substituted in order to decrease overall hydrophobicity. Alignment to CYP52A12 from *C. viswanathii*, also induced upon oleic acid addition [23], and to CYP52A9 from *C. maltosa*, able to terminally oxidize oleic acid [44], was performed (Figure 3.4). Following substitutions were made based on this alignment: F45V, F46T, F54T and F56I. The final polypeptide chains of both CYPtr and CPRtr are shown schematically in Figure 3.5.

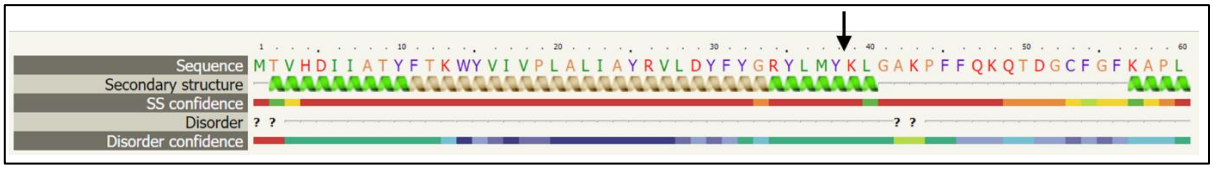

Figure 3.3: N-terminus CYP52A13 – secondary structure and disorder prediction using Phyre2 [317]. The arrow indicates the site of truncation.

Figure 3.4: Alignment of *C. viswanathii* CYP52A13 (tr|Q874J4|Q874J4_CANTR) to *C. maltosa* CYP52A9 (sp|Q12586|CP52I_CANMA) and *C. viswanathii* CYP52A12 (tr|Q874J5|Q874J5_CANTR). Only the N-terminus is shown. The arrows indicate the amino acids that were substituted.

The gene encoding the nontruncated CYP52A13 (CYPnt) was codon optimized, synthesized and cloned in the vector pPICZA in between the restriction sites XhoI and SalI, present in the multiple cloning site of pPICZA, by GenScript. The C-terminal Histag in the vector was excluded by keeping the stop codon. The gene encoding the nontruncated CPR-a (CPRnt) was codon optimized, synthesized and cloned in the vector pAO815 at the EcoRI restriction site, present in the multiple cloning site of pAO816, by GenScript. The codon CUG was modified in the nucleotide sequence as this codon encodes Ser instead of Leu in *C. viswanathii.* The constructs are denoted as CYPnt_ pPICZA and CPRnt_pAO815, respectively. The expression vectors and codon optimized sequences are found in the Addendum. The final polypeptide chains are shown schematically in Figure 3.5.

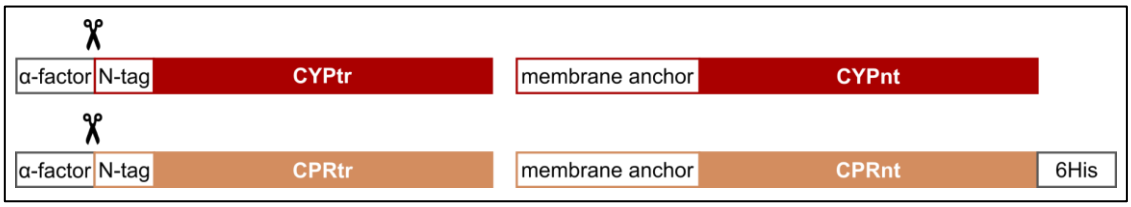

Figure 3.5: Schematic representation of polypeptide chains, as ordered from GenScript. The upper figures show the truncated CYP52A13 and CPR-a. The α-factor secretion signal is present in the vector pPICZαB and the sequence encoding CYPtr and CPRtr were cloned in frame. This secretion signal is processed and cleaved off in the secretory pathway. The N-terminal membrane anchor was removed and an artificial N-tag was added (MGSSHHHHHHSSGLVPRGSHMLDIM). The lower figures show the nontruncated CYP52A13 and CPR-a, where the N-terminal membrane anchor was not removed. In case of CPR-a, the stopcodon was removed in order to add the Histag, present in the pAO815 expression vector.

2.4 Molecular cloning

All restriction enzymes were purchased from NEB. PCR amplification for cloning purposes was performed with the Phusion® High-Fidelity DNA Polymerase from Thermo Scientific. For colony PCR, the Taq DNA Polymerase from Qiagen was used. All primers were synthesized by Integrated DNA Technologies. PCR products were purified using the SpinPrep™ PCR Clean-Up Kit (Novagen). Gel purification was performed with the Wizard® SV Gel and PCR Clean-Up System (Promega). For ligation, the T4 DNA ligase was purchased from Promega.

For molecular cloning creating the CYPtrCPRtr chimeric constructs (Figure 3.6), two different approaches were followed. In a first approach, the CPRtr_ pPICZαB vector was PCR amplified, using the primers ForCPRtr_NdeItoNheI and RevCPRtr_NdeItoNheI (Table 3.1), resulting in a linearized product. The N-terminal tag was excluded and a NheI restriction site was included in the forward primer to replace the NdeI restriction site from this tag-encoding sequence. *cyptr* was PCR amplified from the CYPtr_pPICZαB vector, using the primers ForCYPtr_fusionPCR and RevCYPtr_fusionPCR (Table 3.1). The reverse primer included a sequence encoding a glycine-rich linker. PCR amplified products were purified. Subsequently, a PstI-NheI double digest was performed, after which the digested products were gel purified. The *cyptr* insert was ligated into the linearized vector CPRtr_ pPICZαB. 4 µl of the ligation mixture was added to 80 µl electrocompetent DH5α cells for transformation by electroporation. Transformants were selected on LSLB medium including 50 µg/ml Zeocin™ (InvivoGen). In a second approach, fusion PCR was applied. In a first step, *cyptr* and *cprtr* were PCR amplified from their

respective vectors CYPtr_pPICZαB and CPRtr_pPICZαB, using the primer pairs ForCYPtr_fusionPCR-RevCYPtr_fusionPCR and ForCPRtr_fusionPCR-RevCPRtr_fusionPCR (Table 3.1), respectively. A sequence encoding a glycine-rich linker was included in both the reverse primer for *cyptr* amplification and in the forward primer for *cprtr* amplification (Table 3.1), and the sequence encoding the artificial N-tag was excluded from *cprtr*. The purified PCR products were subsequently fused in a second PCR reaction. 100 ng amplified *cyptr* was added to an equimolar amount of amplified *cprtr*. The fusion PCR reaction existed of two steps. Firstly, ten cycles were executed without primer addition, following a two-step protocol. This was started with an initial denaturation of 30 s at 98 °C. Ten cycles of 10 s denaturation (98 °C) and extension at 72 °C for 30 s/kb (according to the length of the fusion construct), without an annealing step, were performed, finalized by an extension step for 10 min at 72 °C. Secondly, the fusion construct was amplified using the common three-step protocol (25 cycles, 10 s denaturation at 98 °C, annealing for 30 s and extension at 72 °C), including the primers ForCYPtr_fusionPCR and RevCPRtr_fusionPCR (Table 3.1). Both the purified fusion construct and an empty pPICZαB vector were digested with PstI and SalI. After gel purification, the fusion construct was ligated into the digested vector and transformation was performed as described above.

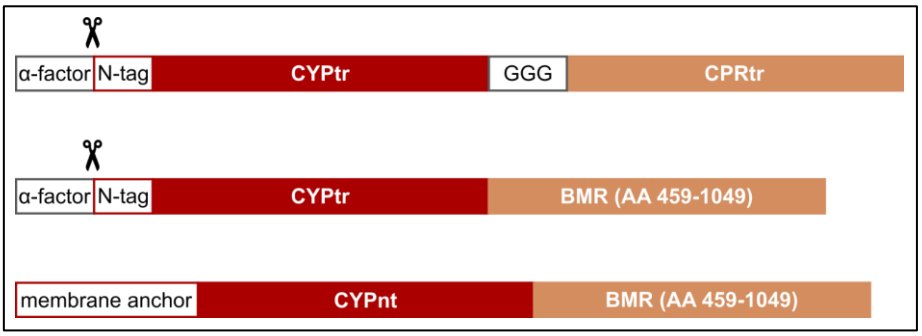

Figure 3.6: Schematic representation of polypeptide chains displaying the different chimeric constructs. (top) The sequence encoding the N-tag was removed from CPRtr and CPRtr was fused C-terminally to CYPtr, linked by three glycine residues. (middle) The reductase domain of CYP102A1 from *Bacillus megaterium* (BMR) was fused N-terminally to the C-terminus of CYPtr. The natural CYP102A1 linker sequence (AA 459-478 [190]) was used to create the chimeric construct. (bottom) BMR, including the natural CYP102A1 linker sequence, was fused N-terminally to the C-terminus of CYPnt.

For both the CYPntBMR chimer as well as the CYPtrBMR chimer (Figure 3.6), Circular Polymerase Extension Cloning (CPEC) [318], was applied. The sequence encoding the reductase domain of CYP102A1, otherwise known as BM3, was amplified from genomic DNA of *Bacillus megaterium* (kindly provided by Prof. Inge Van Bogaert (CSB, Ghent University)) using the primers ForBM3_BMRinclLinker and RevBM3_BMR (Table 3.1), and will be referred to as BMR. The vector, i.e. CYPnt_pPICZA or CYPtr_pPICZαB, respectively, was PCR amplified in order to include the BMR-overlapping sequences, resulting in a linearized plasmid. The respective primer pairs ForCYPnt_CPEC BMR- RevCYPnt_CPEC BMR and ForCYPtr_CPEC BMR- RevCYPtr_CPEC BMR are shown in Table 3.1. The PCR products were gel purified. The *bmr* gene insert was added in an equimolar amount to 100 ng of the respective vector for subsequent CPEC. The reaction was carried out in 25 µl, containing 1X Phusion HF Buffer, 0.4 mM dNTPs, 3 % (v/v) dimethyl sulfoxide (DMSO) and 1 U Phusion® High-Fidelity DNA Polymerase, next to the insert and vector DNA. After an initial denaturation

of 30 s at 98 °C, 15 cycles of 10 s denaturation at 98 °C, 30 s annealing at 55 °C and extension at 72 °C for 15 s/kb (according to the length of the fusion construct), were performed. The reaction was concluded with a final extension step for 10 min at 72 °C. Transformation was performed as described above.

Table 3.1: Primers used for molecular cloning of the different expression constructs. The hyphen separates the nonannealing part of the primer from the sequence annealing to the DNA template. Bold sequences indicate introduced overlapping regions between the two inserts (fusion PCR) or between vector and insert (CPEC). Underlined sequences mark the restriction sites.

Primer	Sequence (5'-3')
ForCPRtr_NdeItoNheI	AGTGGAGCTAGC-CTTGATATTATGTCCCTTAACACCG
RevCPRtr_NdeItoNheI	GTGGTGTGAAGAACCCATAGC
ForCYPtr_fusionPCR	AAGAGAGGCTGAAGCTGC
RevCYPtr_fusionPCR	**ATCCAAGCTAGCGCCACCGCCAGACCTAGG**-ATACATTTCAATATTGGCACCATCG
ForCPRtr_fusionPCR	**CCTAGGTCTGGCGGTGGCGCTAGCTTG**-GATATTATGTCCCTTAACACCGATTCTGG
RevCPRtr_fusionPCR	TCCTGC-GTCGACCTACCAAACATCTTCC
ForCYPnt_CPEC BMR	**GACGTGTGGGCTGGGTAA**-CATCATCATCATCATCATTGAG
RevCYPnt_CPEC BMR	**CTGTTCAGTGCTAGGTGAAGGAAT**-GTACATTTCAATATTAGCACCGTC
ForBM3_BMRinclLinker	ATTCCTTCACCTAGCACTGAACAG
RevBM3_BMR	TTACCCAGCCCACACGTC
ForCYPtr_CPEC BMR	**GACGTGTGGGCTGGGTAA**-CATCATCATCATCATCATTGAG
RevCYPtr_CPEC BMR	**CTGTTCAGTGCTAGGTGAAGGAAT**-ATACATTTCAATATTGGCACC
ForCYPnt_Histag	GTAGGAGCTAG-CCATCATCATCATCATCATTGAG
RevCYPnt_Histag	TCCTAAGCTAGC-GTACATTTCAATATTAGCACCGTC
ForCYPntBMR_Histag	GTAGGAGCTAGC-CATCATCATCATCATCATTGAGTTTTAGCC
RevCYPntBMR_Histag	TCCAACGCTAG-CAGCCCACACGTCTTTTGC
ForCYPnt_KozakSeq	GTAGGACTGCAGGAGATGGGT-ACCGTCCACGACATCATCG
RevCYPnt_KozakSeq	TCCTAACTGCAG-GAGGTACCGATCCGAGACG
ForCYPnt_TrlInCPRnt	GTAGGACTGCAGACCATGGCA-ACCGTCCACGACATCATCG
ForCPRnt_exclHistag	TCCTAACCTA-GGCACCACCACCACCACC
RevCPRnt_exclHistag	GTAGGACCTAGGCTA-CCAAACATCCTCCTGGTATCTATTCTG
For_SDMPmeI	CCAAAACTGACAGTTTA-G-ACGCTGTCTTGGAACCTAATATG
Rev_SDMPmeI	GGTTCCAAGACAGCGT-C-TAAACTGTCAGTTTTGGGCCATTTG
ForCYP_GoldenGate	GTTGAAGCTCTTC-ATCCCCCACACACCATAGC
RevCYP_GoldenGate	TACTAAGCTCTTC-GGATCCGCACAAACGAAGGTC
ForCPRnt_GoldenGate	GTTGAAGCTCTT-CGATCTAACATCCAAAGACGAAAGG
RevCPRnt_GoldenGate	TACTAAGCTCTTC-AGGATCCGCACAAACGAACG

Back-to-back primer pairs were designed for the removal of the stop codon in order to include the Histag at the C-terminus of both CYPnt and CYPntBMR, denoted in Table 3.1 as ForCYPnt_Histag - RevCYPnt_Histag and ForCYPntBMR_Histag - RevCYPntBMR_Histag, respectively. Using these primer pairs, the respective expression vectors were amplified, excluding the stop codon and replacing the SalI restriction site (preceding the Histag) with the NheI restriction site. The PCR product was subsequently digested with NheI, followed by ligation and transformation as described above. The same approach was used for the introduction of two alternative translation initiation sequences in CYPnt_pPICZA. In case of the first alternative sequence, the back-to-back primers ForCYPnt_KozakSeq and RevCYPnt_KozakSeq were used, thereby not only introducing an alternative initiation sequence, but replacing the XhoI restriction site with a PstI restriction site. In case of the second alternative translation initiation sequence, the back-to-back primers ForCYPnt_TrlInCPRnt and RevCYPnt_KozakSeq were used, again replacing the XhoI restriction site with a PstI restriction site. Lastly, this approach was used to exclude the Histag from the CPRnt_pAO815 construct. The primers ForCPRnt_exclHistag and RevCPRnt_exclHistag were designed to include a stopcodon and replace the SalI restriction site (preceding the Histag) with the AvrII restriction site.

To remove the PmeI restriction site from the AOX1 promotor in the expression vector CPRnt_pAO815, a point mutation was performed based on the QuikChange Site-Directed Mutagenesis kit from Agilent with modifications. Primers were designed as described by Zheng et al [319]. Instead of the Pfu DNA polymerase, the Phusion® High-Fidelity DNA polymerase was used. The 50 µl PCR mixture contained 100 ng plasmid DNA, 1X Phusion HF Buffer, 0.2 mM dNTPs, 3 % (v/v) DMSO, 1 U Phusion® High-Fidelity DNA Polymerase and the primer pair For_SDMPmeI - Rev_SDMPmeI (Table 3.1) to a final concentration of 0.5 µM. The same thermocycler program as used in case of CPEC was applied. After PCR, 1 µL DpnI (20 U) was added directly to the PCR reaction mixture and incubated for 1 h at 37 °C in order to digest the original methylated plasmid. 1 µL of this reaction mixture was then added to DH5α chemocompetent cells for subsequent heat shock transformation (30 min incubation on ice after DNA addition, 45 s at 42 °C and 2 min on ice). Transformants were selected on LB medium including 100 µg/ml Cb (Gold Biotechnology).

For the construction of all CYP-CPRnt co-expression plasmids (CYP being either CYPnt or CYPtr) (Figure 3.7), a Golden Gate-inspired approach was used, making use of type IIs restriction enzymes. The *cyp*-containing vector was linearized by PCR using primers including a SapI restriction site (ForCYP_GoldenGate and RevCYP_GoldenGate,Table 3.1). The complete expression cassette for *cprnt* expression was amplified and again primers including the SapI restriction site were used (ForCPRnt_GoldenGate and RevCPRnt_GoldenGate,Table 3.1). After PCR, 1 µL DpnI (20 U) was added and incubated for 1 h at 37 °C. Following purification of the DpnI-treated PCR products, the assembly reaction was carried out in 15 µl, containing 100 ng of the *cyp*-containing vector and an equimolar amount of the *cprnt* expression cassette insert, 1 µl SapI, 1 µl T4 DNA ligase and 1.5 µl T4 DNA ligase buffer. The assembly of the co-expression vector was performed using following thermocycler protocol: 25 cycles

alternating between 37 °C for 2 min and 16 °C for 5 min, followed by a final 5 min incubation at 50 °C and 5 min at 80 °C. The complete assembly reaction was added to DH5α chemocompetent cells for subsequent heat shock transformation (30 min incubation on ice after DNA addition, 90 s at 42 °C and 5 min on ice). Transformants were selected on LSLB medium including 50 µg/ml Zeocin™ (InvivoGen).

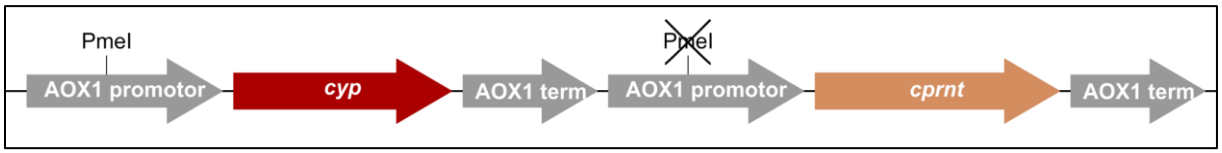

Figure 3.7: Expression cassette for the co-expression of *cyp* (either the truncated or nontruncated variant) and *cprnt* (either including or excluding a Histag). The PmeI restriction site in the AOX1 promotor of *cprnt* was deleted in order to be able to linearize the co-expression vector for efficient integration in the genome of *P. pastoris*.

In all cases, the resulting plasmids were evaluated by colony PCR (primers shown in Table 3.2) and restriction digest.

Table 3.2: Primers used to evaluate the inserted genes

Primer	Sequence (5'-3')
ForCYPtr	GATCTTCCCATCACCATCACC
ForCPRtr	ATGGGTTCTTCACACCACCAC
RevAOX1term	GCAAATGGCATTCTGACATCC
ForCYPnt1	ATCGGTACCTCGAGATGACC
ForCPRnt	GAATTCACCATGGCATTGG
RevCPRnt	CCAAACATCCTCCTGGTATCTATTCTG
ForCYPnt2	ACCGTCCACGACATCATC

2.5 Transformation of constructs in *P. pastoris*

Constructs obtained after molecular cloning were amplified in *E. coli* DH5α. In case of the constructs obtained from GenScript, 10 ng of plasmid DNA was transformed in competent XL1-blue cells (CYPtr_pPICZαB and CPRtr_ pPICZαB) or in competent DH5α cells (CYPnt_pPICZA) by electroporation. Transformed cells were selected on LSLB medium including 50 µg/ml Zeocin™ (InvivoGen). All other constructs were transformed as described in previous section. Plasmid DNA was purified using the NucleoSpin® Plasmid EasyPure (Macherey-Nagel) kit. The plasmid DNA was subsequently linearized with SacI-HF, or PmeI in case of the co-expression constructs, after which the DNA was purified using the SpinPrep™ PCR Clean-Up Kit (Novagen). For transformation by electroporation, typically 5 µl of linearized DNA was added to 80 µl competent *P. pastoris* cells. Transformants were selected on YPD agar including 100 µg/ml Zeocin™ (InvivoGen), incubating at 30 °C for 2 to 5 days. Colony PCR was performed to check the presence of the inserted gene, as described by Lõoke et al [320], using a forward primer, specific for the gene of interest, and a reverse primer annealing to the AOX1 terminator (Table 3.2).

2.6 Recombinant protein production in *P. pastoris*

A preculture was set up by inoculating 10 ml YPD including 100 µg/ml Zeocin™ (InvivoGen) in a 50 ml birmboin tube with a colony, or alternatively, with 1 ml from a glycerol stock. The preculture was incubated overnight at 28 °C and 250 rpm. *P. pastoris* was first grown in BMGY medium, inoculated with preculture (1/50), for 48 h at 28 °C and 250 rpm. After the growth phase, cells were induced by a medium change to BMMY. To this end, the cells were washed with BMMY and the supernatant was thrown away to remove residual glycerol. The cell pellet was subsequently resuspended in BMMY. Protein production was performed at 28 °C and 250 rpm for 48 h, unless otherwise specified. The expression was maintained by methanol addition every 12 h to a final concentration of 1 %, unless otherwise specified.

The small-scale expression tests were carried out in 50 ml birmboin tubes, inoculated with a 10 ml culture. 100 ml cultures in a 1 l shake flask were set up in case of microsome isolation for subsequent enzymatic assays. For enzyme collection and Immobilized Metal Affinity Chromatography (IMAC) purification from the medium, 500 ml cultures were grown and induced in 3 l baffled shake flasks. The milli-scale expression tests were performed in 24-well deep-well plates with round bottoms, inoculated with 2 ml culture.

2.7 Cell lysis and microsome isolation

For analysis of the lysate or microsomes after a milli-scale or small-scale expression test, cells from a 1 ml sample were collected and washed in 50 mM phosphate buffer, pH 7.4, 5 % glycerol. Subsequently the cell pellet was resuspended in 100 µl breaking buffer (50 mM phosphate buffer, pH 7.4, 5 % glycerol, 1 mM ethylenediaminetetraacetic acid (EDTA), protease inhibitor cocktail). An equal volume of acid washed glass beads was added and cells were disrupted by vortexing the suspension for 30 s, alternated by a 30 s incubation on ice to prevent cells from heating up. This cycle was repeated eight times. Cell debris and glass beads were pelleted by centrifugation for 10 min at 12 000 g and 4 °C. The supernatant was either collected for lysate analysis, or further ultracentrifuged for 1 h at 100 000 g and 4 °C for microsome collection. The collected microsomes were washed and resuspended either in resuspension buffer (50 mM phosphate buffer, pH 7.4, 20 % glycerol, 1 mM EDTA, 1 mM dithiothreitol (DTT) and protease inhibitor cocktail) or in Laemmli buffer. For CYP quantification by Multiple Reaction Monitoring (MRM), microsomes were resuspended in 50 mM NH_4HCO_3 containing 0.1 % (w/v) RapiGest™ SF (Waters). Alternatively, the resuspension buffer was used instead of the breaking buffer for cell disruption. In case of microsome collection for enzymatic assays, the same protocol as described above was applied, except that cells from a 100 ml culture were collected, washed and resuspended in breaking buffer to a final OD_{600} of ca. 100. Enzymatic cell lysis was performed as described by Pompon et al [321].

2.8 Protein concentration from the medium

In order to analyse medium samples after a small-scale or milli-scale expression test, a 1 ml sample from the medium was taken for subsequent protein concentration by DOC-TCA precipitation. Firstly, 100 µl sodium deoxycholate (DOC) (0.5 % w/v) was added to this sample and incubated on ice for 10 min. Secondly, 110 µl trichloroacetic acid (TCA) (100 % w/v) was added and the sample was further incubated on ice for 20 min. Alternatively, the samples were incubated overnight on ice after TCA addition. The precipitated protein was subsequently pelleted by centrifugation at 16 000 g and 4 °C for 30 min. The pellet was washed two times with 1 ml ice-cold acetone (Biosolve) and one time with ice-cold 70 % ethanol (VWR). The pellet was dried and resuspended in phosphate buffered saline (PBS), or directly in Laemmli buffer for further analysis. Alternatively, the medium sample was concentrated by ultrafiltration using the Amicon® Ultra-0.5 centrifugal filter devices (Merck) with a 10 kDa cutoff.

Protein concentration from the medium after CPRtr production in a 500 ml culture was performed by ammonium sulphate precipitation. $(NH_4)_2SO_4$ was added to a saturation of 75 % at 0 °C and the medium was stirred gently overnight at 0 °C for complete precipitation. The next day, the precipitated protein was collected by centrifugation at 10 000 g and 0 °C for 30 min. The supernatant was collected and $(NH_4)_2SO_4$ was added to a saturation of 90 % at 0 °C. The supernatant was stirred gently at 0 °C for 1 h and the precipitated was collected as before. Both pellets were resuspended in 50 ml 50 mM Tris HCl (pH 7.5), including 20 % glycerol, 0.5 µM FAD and FMN and 0.5 mM DTT. Desalting was performed using Amicon® Ultra-15 centrifugal filter devices (Merck) with a 10 kDa cutoff.

2.9 SDS-PAGE and western blot

Protein samples were mixed with Laemmli buffer and the proteins were separated by SDS-PAGE using a 12 % polyacrylamide gel and a Tris-glycine running buffer (Bio-Rad). The Precision Plus Protein™ Unstained Standard (Bio-Rad) was used as a molecular weight marker. Gels were stained with Coomassie brilliant blue G.

In case of western blot analysis, SDS-PAGE was performed as described above, using the Precision Plus Protein™ Dual Color Standard (Bio-Rad). After protein separation was completed, the proteins were transferred to a nitrocellulose blotting membrane. The membranes were incubated with a 5 % (w/v) nonfat dry milk (Bio-Rad) solution in PBS and washed with PBS-T (PBS including 0.1 % (v/v) Tween-20) before incubation with a horseradish peroxidase (HRP)-conjugated 6x-His epitope tag monoclonal antibody, diluted in a 1 % (w/v) nonfat dry milk (Bio-Rad) solution in PBS-T. For the C-terminal Histag, the anti-His(C-term)-HRP antibody (R931-25, Invitrogen) was used, at a dilution of 1:10 000. In case of an N-terminal Histag, the anti-His-HRP antibody (MA1-21315-HRP, Invitrogen) was used, at a dilution of 1:5000. After washing with PBS-T, proteins were visualized by chemiluminescence.

2.10 IMAC purification

The medium was collected after protein production in a 500 ml culture. Reduced glutathione was added to a final concentration of 0.01 % (w/v) and $MgSO_4$.7H2O was added to a final concentration of 2 mM, after which the pH was adjusted to 7.5. The medium was subsequently filtered using a Steritop® Filter Unit (EMD Millipore) or VacuCap® (VWR) with a pore size of 0.22 µm. IMAC was performed using a HisTrap™ HP column (5 ml, Cytiva), mounted on the ÄKTA purifier (Cytiva). The pretreated sample was loaded on the IMAC column at a flow rate of 5 ml/min after equilibration with binding buffer (50 mM sodium phosphate, pH 7.5, 0.5 M NaCl). Following sample loading, the column was washed with binding buffer and when the absorbance reached a steady baseline, the flow rate was reduced to 1 ml/min for subsequent elution using the binding buffer including increasing imidazole concentrations. A stepwise elution was performed, first using 10 mM imidazole and secondly with 100 mM imidazole. A third step with 250 mM imidazole was included in the first purifications and was discarded if no protein of interest eluted in this step. The last elution step was performed using 500 mM imidazole. Immediately after elution, a buffer exchange was performed with 25 mM Tris-HCl, pH 7.5, 150 mM NaCl, unless otherwise specified, using the Amicon® Ultra-15 centrifugal filter devices (Merck) with a 10 kDa cutoff. Alternatively, $(NH_4)_2SO_4$ precipitation was performed as described in 2.8.

2.11 Determination of protein concentration

Protein concentrations were determined with the Thermo Scientific™ Coomassie (Bradford) Protein Assay Kit. The measurement was done using the Bio-Rad Microplate Reader model 680.

2.12 Trypsin digestion and MALDI-TOF MS

Gel bands of interest were cut out after separation by SDS-PAGE and visualization of the proteins with Coomassie brilliant blue G. After complete removal of the Coomassie stain using 50 % acetonitrile (ACN) (Biosolve) in 200 mM NH_4HCO_3 at 30 °C, a reduction step with 10 mM DTT in 100 mM NH_4HCO_3 for 1 h at 56 °C and an alkylation step with 55 mM iodoacetamide (IAA) in 100 mM NH_4HCO_3 for 45 min at room temperature in the dark was performed. Gel bands were subsequently washed with 100 mM NH_4HCO_3 and dehydrated with 100 % ACN (Biosolve) (two times). Following complete dehydration of the gel bands, 0.02 µg modified trypsin (Promega) in 10 µl 50 mM NH_4HCO_3 was added and samples were incubated on ice for 45 min. 40 µl 50 mM NH_4HCO_3 was added and the proteins were digested overnight at 37 °C. The next day, peptides were extracted twice with a 60 % ACN (Biosolve) solution containing 0.1 % formic acid. The extracts were pooled and dried under vacuum.

The dried peptides were then resuspended in 10 µl of a 50 % ACN (BioSolve) solution containing 0.1 % formic acid. 1 µl of resuspended peptides, mixed with a saturated α-cyano-4-hydroxycinnamic acid solution in a 1:1 ratio was spotted onto an Opti-TOF 384 Well MALDI Plate Insert for Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) analysis with the MALDI TOF/TOF 4800 Plus (ABSciex). Identification was done with Mascot using the *P. pastoris* NRRL-Y-11430 protein database, downloaded from Uniprot.

2.13 Enzymatic assays

2.13.1 Oleic acid conversion

The *in vitro* reaction using isolated microsomes was carried out in a 1 ml reaction mixture, containing; 200 mM potassium phosphate, pH 7.4, 1 mg/ml microsomal protein, 0.5 mM DTT, 0.1 mM oleic acid (from a 100 mM stock solution in DMSO) and 1 mM NADPH. The reaction was incubated at 35 °C and 150 rpm for 1 h. The conversion was stopped by adding 500 µl 6 M HCl and extraction was performed with and equal volume of diethyl ether (three times). The diethyl ether extracts were pooled, evaporated under vacuum and stored at – 20 °C until GC-MS analysis.

The *in vitro* whole-cell bioconversion assay was performed as follows; expression was carried out in 2 ml cultures in a 24-well deep-well plate as described before. After a 48 h expression phase, cells were washed and resuspended in 100 mM potassium phosphate buffer, pH 7.4. Oleic acid was added to a final concentration of 0.1 mM, 1 mM or 10 mM from a 350 mM stock in DMSO. The deep-well plate was incubated at 28 °C and 320 rpm for 8 h. The reaction was stopped by HCl addition and the formed product was extracted from the culture supernatant with diethyl ether as described above.

For the *in vivo* bioconversion assay, cells were grown in 50 ml birmboin tubes. After a 48 h growth phase, the cells were collected and resuspended in 5 ml BMMY. Oleic acid was added to a final concentration of 0.2 % (v/v) and incubated for 48 h. A 500 ml sample of the culture was taken every 12 h, extracted with diethyl ether and evaporated under vacuum as described above.

For the optimization of the *in vitro* assay, using microsomes containing CYPnt, five different assays were applied, selected from different sources assaying ω-hydroxylation and are shown in

Table 3.3: Five different enzymatic assays, selected from literature and all assaying ω-hydroxylation, applied for the CYPnt-CPRnt-containing microsomes.

	Assay 1 [322]	**Assay 2** [26]	**Assay 3** [24]	**Assay 4** [79]	**Assay 5** [82]
Microsomal protein	5 mg	50 µg	0.5 mg	0.5 mg	0.5 mg
Phosphate buffer, pH 7.4	50 mM	100 mM	200 mM	200 mM	50 mM
Oleic acid	1 mM	5 µM	15 µM	50 µM	1 mM
NADPH	10 mM	4.5 mM	0.5 mM	0.5 mM	1 mM
DTT			0.5 mM	0.5 mM	
$MgCl_2$					1 mM
G6P			3 mM	3 mM	5 mM
G6PDH			0.5 U	0.5 U	12 U
Volume	**1 mL**	**250 µL**	**250 µL**	**1 mL**	**500 µL**
30 °C, 500 rpm	**10 min**	**30 min**	**1 h**	**250 µL/h**	**4 h**

The assays using microsomes collected after expression of the fusion enzyme CYPtrCPRtr, followed by GC-MS analysis of silylated end products, were performed in parallel with the surface display whole-cell bioconversion assays, where the 1 ml reaction mixture contained 100 mM potassium phosphate, pH 7.4, 0.1 mM oleic acid (from a 10 mM stock in DMSO), 1 U glucose-6-phosphate dehydrogenase, 5 mM glucose-6-phosphate and 1 mM NADPH. The reaction was carried out for 3 h at 37 °C and 500 rpm, after which the conversion was quenched with HCl and an extraction with diethyl ether was performed as described above.

2.13.2 NADPH oxidation assay

On the one hand, an NADPH oxidation assay was carried out in a volume of 250 µl in a microtiter plate using the Bio-Rad Microplate Reader model 680. The reaction mixture was composed of 200 mM potassium phosphate, pH 7.4, 1 mg/ml microsomal protein, 0.5 mM NADPH and four different concentrations of oleic acid (0, 0.025, 0.050, 0.1 mM) were added. The reaction was initiated by NADPH addition and the absorbance was measured at 340 nm.

On the other hand, the assay was carried out in a volume of 1 ml in a UV-VIS cuvette using the spectrophotometer Genesys 10S UV-VIS from Thermo Scientific. The volume of this assay was increased to 1 ml in order to decrease the DMSO concentration in the final reaction. The reaction mixture was composed of 200 mM potassium phosphate, pH 7.4, 1 mg/ml microsomal protein, 0.2 mM oleic acid (from a 100 mM stock in DMSO) and 2 mM NADPH. The reaction was initiated by NADPH addition and carried out for 1 h at 28 °C. The absorbance difference between a reaction mixture including and excluding oleic acid was monitored at 340 nm.

2.13.3 Cytochrome c reductase activity assay

A 1 ml reaction was carried out at 28 °C (or at room temperature in case of CPRnt co-expression) for 5 min and the reduction of cytochrome c was followed at 550 nm, using the spectrophotometer Genesys 10S UV-VIS from Thermo Scientific. The 1 ml reaction mixture consisted of 300 mM potassium phosphate, pH 7.7, 0.04 mM cytochrome c (from equine heart), 0.1 mg microsomal protein or 50 µl CPRtr and 0.1 mM NADPH. The reaction was started by the addition of NADPH.

2.14 GC-MS analysis

For methylation, the evaporated extract was dissolved in 500 µl 2 M H_2SO_4 in methanol. An equal amount of hexane was added and the reaction was shaken for 20 min. The reaction was subsequently centrifuged and the upper hexane layer was isolated for Gas Chromatography (GC)-MS analysis as described by Bauwelinck et al [39].

For silylation, the evaporated extract was dissolved in 50 µl 1 % trimethylchlorosilane (TMCS) in N,O-bis(trimethylsilyl) trifluoroacetamide (BSTFA) and incubated for 30 min at 75 °C. After a centrifugation step at 10 000 g for 10 min, the supernatant was transferred to a glass vial for GC-MS analysis. The GC-MS system used was the 7890-5975C from Agilent in combination with the HP-5MS column (5% phenyl methyl polysiloxane, 30m×0.25mm ID, 0.25µm) from Agilent. The temperature was maintained at 180 °C for 1 min, subsequently raised to 300 °C at 8 °C/min and then held isotherm for 5 min.

A 1,18-octadecendioic acid standard was produced in house using the *C. viswanathii* strain ATCC 20962, as described in [323].

2.15 CYP quantification by multiple reaction monitoring

To 10 µg microsomal protein, 0.04 µg bovine serum albumin (BSA) (Thermo Scientific) was added as an internal standard and 50 mM NH_4HCO_3 containing 0.1 % (w/v) RapiGest™ SF (Waters) was added to a final volume of 50 µl. This was heated to 80 °C for 10 min. Subsequently, a reduction step was performed with 5 µl 50 mM DTT in 50 mM NH_4HCO_3 at 60 °C for 10 min, followed by an alkylation step with 5 µl 150 µM IAA in 50 mM NH_4HCO_3 at room temperature in the dark for 20 min. After 0.2 µg modified trypsin (Promega) addition, the solution was incubated overnight at 37 °C for digestion. The next day, the samples were acidified with 1 µl 8 % formic acid and incubated for 30 min at 37 °C. These were then centrifuged for 30 min at 16 000 g and the supernatant was filtered with a 0.22 µm filter (Corning® Costar® Spin-X® Plastic Centrifuge Tube Filter). Drying of the filtrate was performed under vacuum, after which the dried peptides were redissolved in 100 µl 0.1 % formic acid and desalted using OMIX C18 pipette tips (Agilent). The desalted peptides again were dried under vacuum and dissolved in 0.1 % formic acid, 3 % ACN and 10 fmol/ml phosphorylase B (Hi3 Phos B Standard, Waters).

For CYP quantification by MRM, the following Liquid Chromatography (LC)-MS/MS system was used, i.e. a Waters NanoAcquity M-Class UPLC and an IonKey source connected to a Waters Xevo TQ-S triple quadrupole mass spectrometer. The IonKey source contained a 150 µm x 100 mm, 1.8 µm HSS T3, iKey separation device. Samples were injected (5 µl, i.e. 0.5 µg) and trapped on a 300 µm x 50 mm, 5 µm, 100 Å Acquity UPLC M-Class Symmetry C18 Trap Colum (Waters). Subsequently the peptides were separated on the iKey. The following mobile phase was used; solution A consisted of 0.1 % formic acid (Biosolve) and solution B of 0.1 % formic acid in ACN (Biosolve). The washing step on the trap column was performed for 2 min with 3 % B at a flow rate of 15 µl/min. Separation was performed in 10 min at a flow rate of 2 µl/min using a linear ACN gradient (3-50 %). The auto-sampler was washed with 0.1 % formic acid as the strong solution and 0.1 % formic acid in ACN/water/isopropanol (50:25:25, v/v/v) as the weak solution. Following separation, quantification of the peptides was performed in positive ion mode (ESI+) with the Waters Xevo TQ-S mass spectrometer. The MRM mode with transitions of selected precursor ions at a set cone voltage and different collision energies and dwell times for each precursor, was used for detection (see Addendum). These precursor ions were selected based on predictions made in Skyline. With the online tool Unipept [324] it was checked whether peptides were unique. For electrospray ionization (ESI)-MS/MS, capillary voltage was set at 3.5 kV, the cone voltage at 35 V and the source temperature at 120 °C. In the collision cell, argon was introduced at a flow rate of 0.15 ml/min. MassLynx 4.1 and Skyline software were used for data acquisition and processing, respectively. Savitsky-Golay Smoothing was applied, the total area under the curve (AUC) for each peptide was calculated and normalized to BSA.

2.16 Statistical analysis

The GraphPad Prism software was used for all statistics.

3 Results

3.1 Recombinant production of a truncated CYP52A13 in *Pichia pastoris*

The gene encoding the truncated CYP52A13 as described in material and methods (see 2.3), was codon optimized and cloned in the expression vector pPICZαB in frame with the α-factor secretion signal. Three colonies were picked upon selection with Zeocin™ after transformation of the linearized plasmid in *P. pastoris* and a small-scale expression test was performed. A secretion signal was provided thus CYPtr is expected in the medium. Medium samples were taken, concentrated and analysed by western blot (Figure 3.8A). In parallel, the cells were lysed and a sample of the lysate was taken along (Figure 3.8B). It was observed that CYPtr (58 kDa) was successfully produced in *P. pastoris*. Unfortunately, the enzyme was not secreted to the medium and was only present in the lysate samples. The protein bands also appeared at a higher molecular weight (MW) than expected, suggesting that the secretion signal of 9.3 kDa was not processed. It is known that apart from the N-terminal helix, also the F-G loop is closely interacting with the ER membrane. Therefore, microsomes were isolated from the cell lysate to investigate whether the CYPtr remains bound to these membranes (Figure 3.8C). Indeed, this was the case so the designed truncations appeared insufficient for complete solubilization.

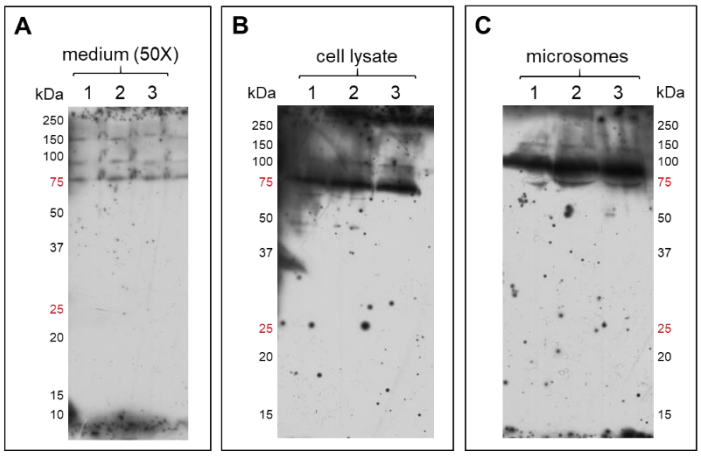

Figure 3.8: Western blot after small-scale expression test of three colonies picked after *cyptr* transformation in *P. pastoris*. A: concentrated medium samples. B: cell lysate. C: isolated microsomes. CYPtr (MW 58 kDa) was not secreted but remained bound to the microsomal membrane. The higher MW indicates that the secretion signal (9.3 kDa) was not processed.

To verify whether this microsome-bound CYPtr was active, an enzymatic assay was performed using the microsomes isolated from a 100 ml culture of CYPtr producing *P. pastoris*. This host possesses an endogenous ER-bound CPR and it was hypothesized that this CPR could serve as the redox partner to donate electrons required for the FA hydroxylation and subsequent overoxidation reaction. Oleic acid was chosen as the substrate, because Eschenfeldt et al showed that CYP52A13 showed the highest activity towards this FA [24]. On the one hand, an assay was performed for which the products were extracted with diethyl ether upon acidification, methylated and analysed using GC-MS, in search for the desired 1,18-octadecenedioic acid end product. On the other hand, NADPH oxidation was measured spectrophotometrically after addition of an increasing oleic acid concentration. It is known that uncoupling reactions can occur during CYP catalysis through three different reactions, i.e. the

autooxidation shunt, the peroxide shunt and the oxidase shunt, where NADPH is consumed without product formation [172], [173]. Measuring the NADPH oxidation thus also takes into account possible uncoupling through one of these shunts. In the GC chromatogram (Figure 3.9), only the substrate oleic acid was identified and the desired DCA was not found (expected retention time, or RT, of 27.45 min). According to the NADPH oxidation assay (Figure 3.10), NADPH concentration was not decreasing at a higher rate upon increasing the oleic acid addition. NADPH was thus not consumed by the recombinantly produced CYPtr. The observed (weak) cofactor oxidation is thus likely due to the action of endogenous microsomal enzymes and/or occurs aspecifically as NADPH is unstable. Two possible reasons for the lack of activity were investigated further; (1) can the endogenous *P. pastoris* CPR indeed act as an alternative redox partner, (2) are the truncations and/or presence of the α-factor secretion signal leading to inactivity?

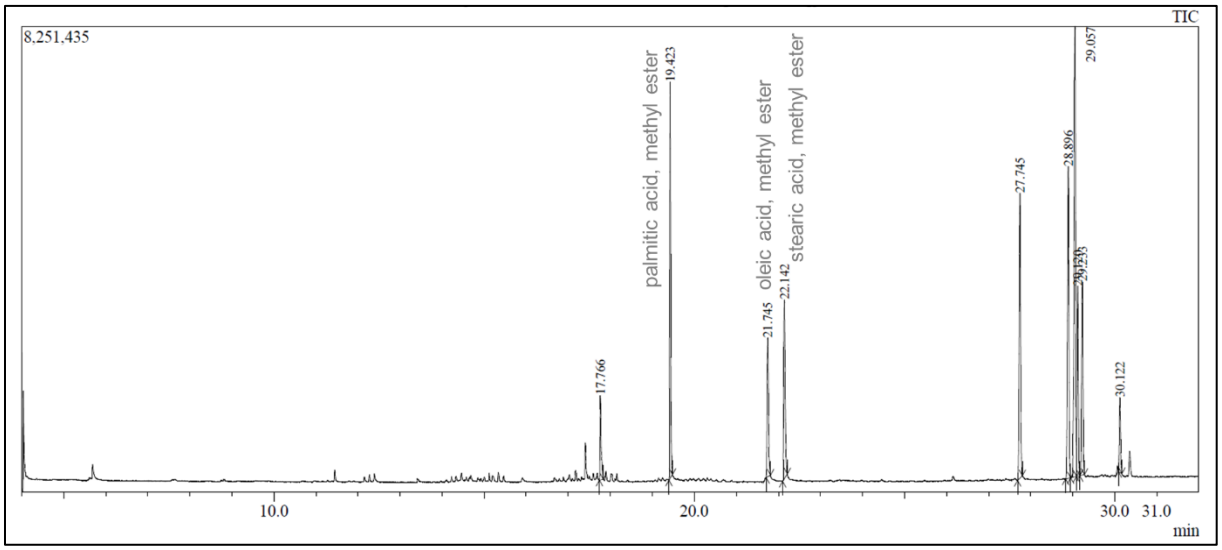

Figure 3.9: GC chromatogram of methylated extracts after microsome activity assay with CYPtr. No conversion of oleic acid to the corresponding DCA (RT = 27.45 min) was observed.

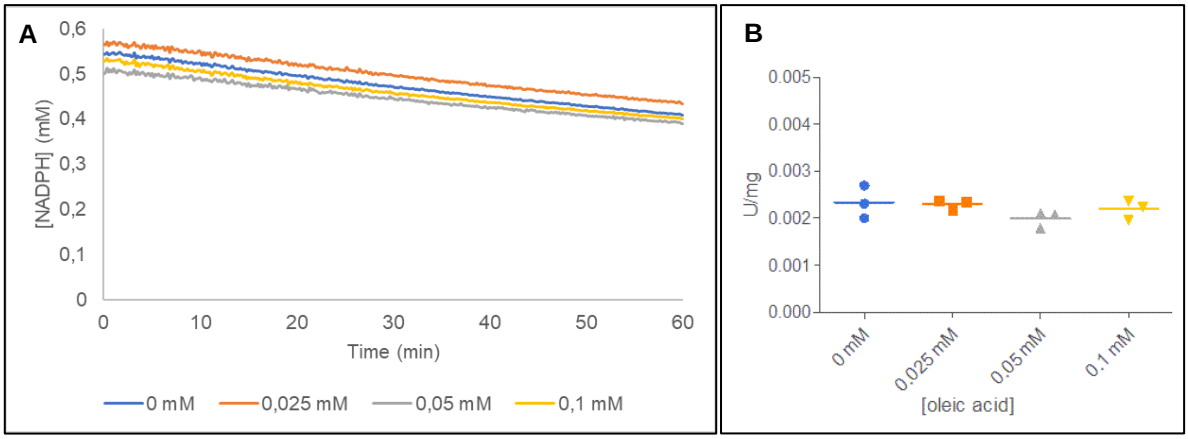

Figure 3.10: NADPH oxidation assay CYPtr-containing microsomes. A: The average of three technical replicates is shown. B: The specific activities did not significantly differ from each other, as one-way ANOVA analysis returns a p-value of 0.3206. Increasing the oleic acid substrate concentration did thus not increase the NADPH consumption in CYPtr-containing microsomes.

3.2 Is the native redox partner CPR-a required for CYP52A13 activity?

It was hypothesized that the endogenous CPR from *P. pastoris* was not compatible with CYPtr or it is produced at too low levels to support the recombinant CYPtr. Indeed, it has been documented that switching to a heterologous redox partner is not without consequence. It can markedly influence the activity and selectivity. Even other functionalities were obtained as was reported by Zhang et al [267]. Therefore, the native CPR-a redox partner was recombinantly produced in *P. pastoris.* A truncated form was produced in order to purify this enzyme for functional assays. A small-scale expression test was performed, taking along three colonies, picked upon transformation of *P. pastoris* with the linearized plasmid. Due to the presence of the α-factor secretion signal and the truncation of the N-terminal anchor, the enzyme is expected in the medium. A medium sample was taken and concentrated. Indeed CPRtr was secreted, shown by western blot analysis (Figure 3.11). Cells were lysed and the lysate was taken along in parallel. A large amount of CPRtr appeared to remain intracellularly and was degraded.

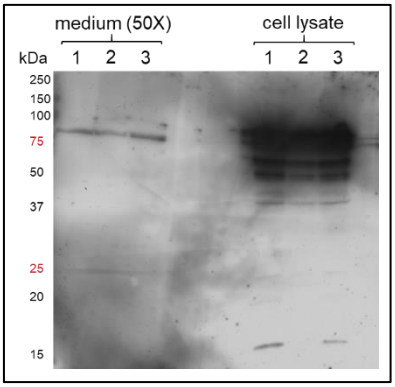

Figure 3.11: Western blot after small-scale expression test of three colonies picked after *cprtr* transformation in *P. pastoris*. CPRtr (MW 76 kDa) was successfully produced and secreted to the medium. However, a large amount of degradation was observed in the lysate.

Part of the recombinantly produced CPRtr was successfully secreted so a first upscaling was done to a 500 ml culture in a 3 l baffled shake flask. The medium was collected and purified by IMAC purification (Figure 3.12). A stepwise elution was performed, and protein eluted mainly at a concentration of 100 mM imidazole. Analysing the collected fractions by SDS-PAGE and western blot (Figure 3.13A and B (1), respectively), an intense band around the MW of 76 kDa (the expected MW of CPRtr) was seen. The enzyme was thus successfully purified from the medium and was already fairly pure, based on the SDS-PAGE gel. Both the gel and blot show a second band at slightly higher MW. Three other expression conditions were tested in parallel. On the one hand, the cultures were incubated at two different temperatures during expression. On the other hand, induction was performed using two different methanol concentrations, i.e. adding methanol every 12 h to a final concentration of 1 % and adding methanol every 24 h to a final concentration of 0.5 %. These two different bands around 76 kDa were consistently observed. As CPRtr contains glycosylation sites, this slightly higher band might indicate a low amount of glycosylation. An alternative explanation is the incomplete processing of the secretion signal. No increased secretion was observed when lowering the temperature to 16 °C during expression and/or decreasing the methanol concentration compared to expression at 28 °C and inducing with 1 % methanol every 12 h (Figure 3.13A and B). Analysing the lysate

showed that altering the expression conditions did not decrease the intracellular degradation either (Figure 3.13C).

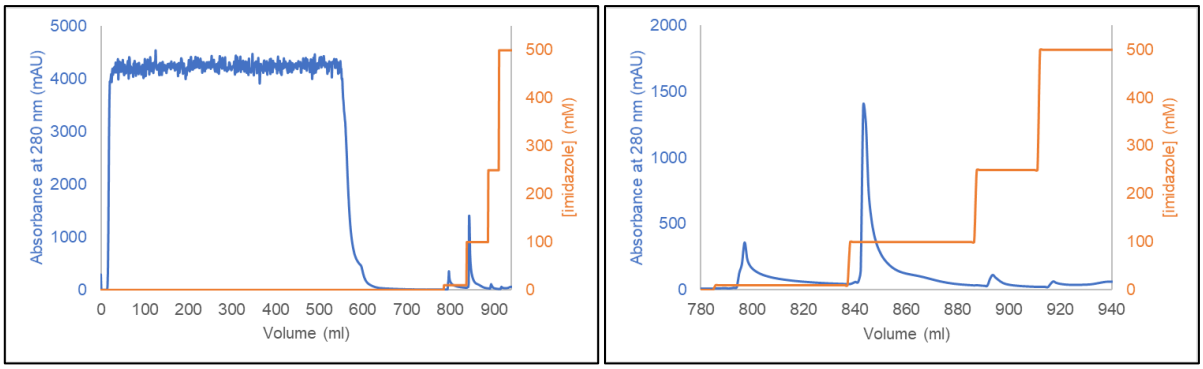

Figure 3.12: IMAC chromatogram of CPRtr purification from the medium after expression in a 500 ml culture. The right chromatogram shows the elution profile in more detail.

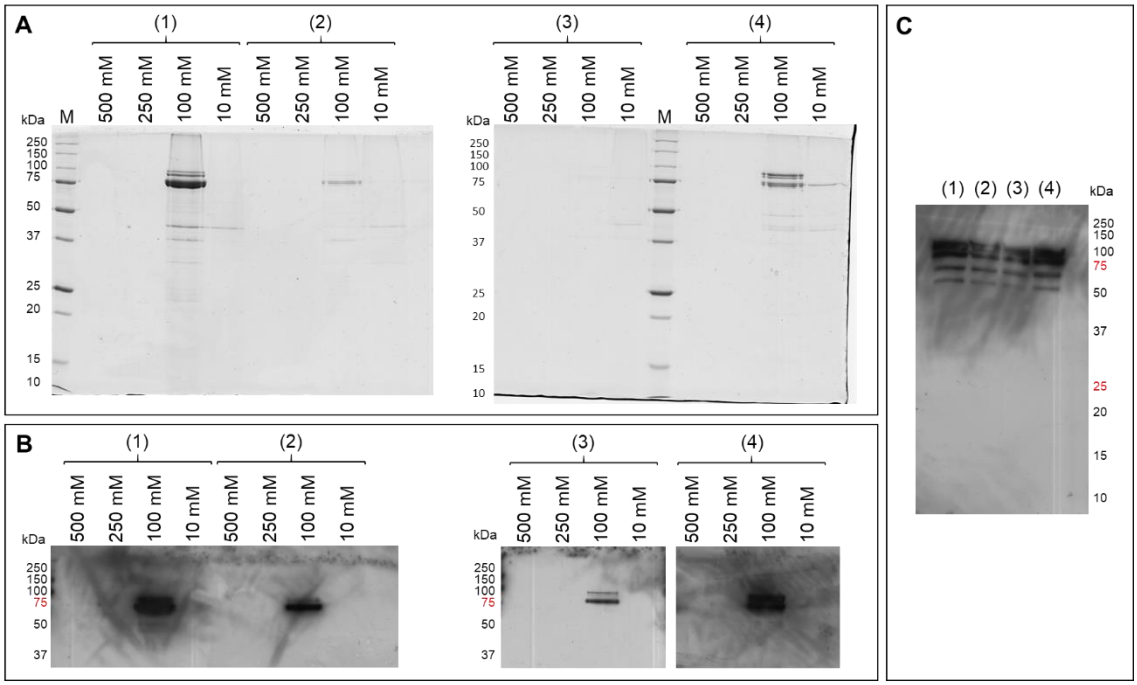

Figure 3.13: A: SDS-PAGE analysis of IMAC elution peaks, desalted and concentrated after CPRtr (76 kDa) purification. B: Western blot analysis of IMAC elution peaks, desalted and concentrated after CPRtr purification. C: Western blot of lysate samples. Four different conditions were tested for the CPRtr expression phase. (1) 28 °C, 1 % MeOH every 12 h, (2) 28 °C, 0.5 % MeOH every 24 h, (3) 16 °C, 1 % MeOH every 12 h, (4) 16 °C, 0.5 % MeOH every 24 h.

In order to improve secretion to the medium and eliminate intracellular degradation, two alternative strains to the wild type *P. pastoris* NRRL-Y-11430 were transformed with linearized CPRtr_pPICZαB. Both alternative strains co-express *HAC1*, encoding the transcription factor Hac1p, upon methanol induction. The transcription factor Hac1p is a central regulator in the unfolded protein response, normally activated by the accumulation of unfolded and misfolded proteins in the ER. This transcription factor activates the expression of chaperones, foldases and proteins involved in lipid metabolism. Guerfal et al showed that inducible co-expression of *HAC1* in *P. pastoris* can lead to an increased recombinant protein yield, either secreted or surface displayed. Moreover, when producing the membrane-bound G-protein coupled receptor, adenosine A2A receptor, provided with an α-factor secretion signal, *HAC1*

co-expression aided in the processing of this secretion signal [315]. In one of the two *HAC1* co-expression strains, *pep4* was knocked out. This gene encodes proteinase A, a vacuolar aspartyl protease, activating other vacuolar proteases such as carboxypeptidase and proteinase B [311], [314]. Eliminating this gene thus significantly reduces intracellular proteolysis.

Subsequent to transformation, a small-scale expression test was performed including several colonies for the two strains (seven in case of the *HAC1* co-expressing strain, eight in case of the *pep4* knock-out strain). Similar as before, both medium and lysate samples were taken for western blot analysis. Not any colony of the *pep4* knock-out strain produced CPRtr and is therefore not further discussed. In case of the *HAC1* co-expressing strain, only one colony was found to produce CPRtr. To concentrate collected medium samples, two methods were applied. Part of the sample was concentrated 25X using an Amicon® ultra-0.5 centrifugal filter device with a 10 kDa cutoff. Alternatively, the sample was concentrated 50X by DOC-TCA precipitation and resuspended in Laemmli buffer. The western blot analysis in Figure 3.14A revealed a bright signal for the latter sample only. This indicates that using the centrifugal filter units could lead to loss of protein. Further investigation of the western blot shows that the lysate contains bands at a lower MW, indicating that the intracellular degradation is not solved by co-expressing *HAC1*. To confirm if *HAC1* indeed was co-expressed, medium samples of both the wild-type strain and the *HAC1* co-expressing strain, both producing CPRtr, were analysed by SDS-PAGE (Figure 3.14B). Guerfal et al reported that upon co-expressing *HAC1*, the two ER HDEL-containing chaperones Kar2p and Pdip, were present in the medium in an increased amount [315]. The medium of the *HAC1* co-expressing strain showed a higher protein content in general and indeed showed two bands with a higher relative intensity, coinciding with the MW of Kar2p (74kDa) and Pdip (58 kDa). Using MALDI-TOF MS, it was confirmed that these two bands indeed contained the aforementioned proteins (data not shown). CPRtr on the other hand, could not be detected on this gel. Is the signal at 76 kDa obscured by the Kar2p signal or is the yield too low to detect on gel? In case of the wild type strain, CPRtr could not be identified either.

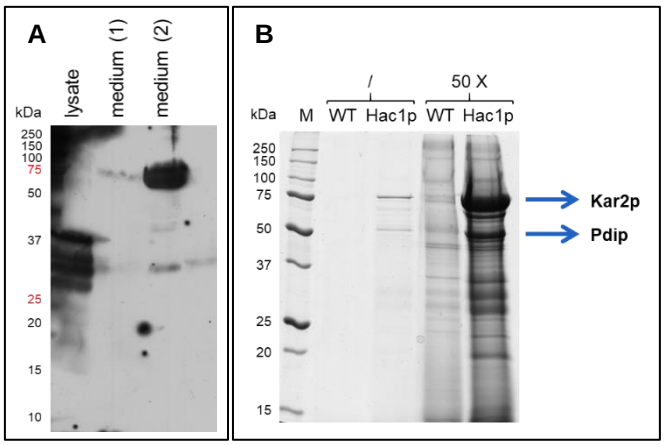

Figure 3.14: Analysis after small-scale expression test of CPRtr (76 kDa) in the *HAC1* co-expressing strain. A: Western blot analysis. (1) Medium sample concentrated 25X using an Amicon® ultra-0.5 centrifugal filter device with a 10 kDa cut-off. (2) Medium sample concentrated 50X by DOC-TCA precipitation. B: SDS-PAGE analysis of concentrated medium sample, collected after CPRtr expression in wild-type *P. pastoris* (WT) and in the *HAC1* co-expressing strain (Hac1p). The intense bands around 75 kDa and 50 kDa were identified as Kar2p and Pdip, respectively, by MALDI-TOF MS.

Following the small-scale expression test, a 500 mL expression test in a 3 l baffled shake flask was performed to see if *HAC1* co-expression led to an increased yield. The medium was collected and IMAC purified (Figure 3.15). Based on the height of the 100 mM imidazole elution peak, no yield increase was obtained using this alternative strain. It must be noted that the filter of the ÄKTA Purifier clogged during sample loading, indicating that precipitation occurred. After buffer exchange, samples of the four elution peaks were analysed by SDS-PAGE (Figure 3.16). Only a very light band around 76 kDa was observed, together with some lower MW bands in the 100 mM imidazole elution peak. The protein thus seemed to have degraded or precipitated during the buffer exchange.

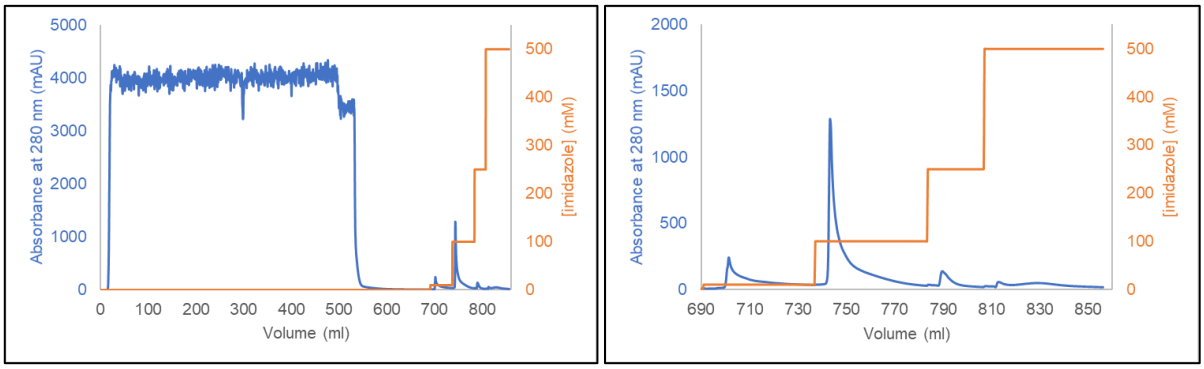

Figure 3.15: IMAC chromatogram of CPRtr purification from the medium after expression in the *HAC1* co-expressing strain in a 500 ml culture. The right chromatogram shows the elution profile in more detail.

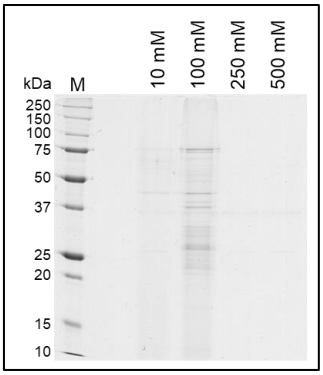

Figure 3.16: SDS-PAGE analysis of IMAC elution peaks, desalted and concentrated after CPRtr (76 kDa) purification Figure 3.15.

The problem of protein precipitation during sample loading on the IMAC column was hypothesized to be a result of CPRtr instability. Therefore, the 500 ml expression test was repeated, medium was collected and IMAC purification was performed, now including 5 % glycerol in the IMAC buffers. The pH adjustment from pH 6 to pH 7.5 in preparation of IMAC was executed more gently using a 1 M K_2HPO_4-solution instead of KOH pellets. No clogging of the filter occurred during sample loading and based on the height of the 100 mM imidazole elution peak in Figure 3.17, the protein yield was significantly improved. Adjusting the pH in a more gently fashion, together with glycerol addition thus seemed to markedly increase the stability during IMAC. Subsequent to IMAC, a buffer exchange was performed as before, now using an adjusted storage buffer, i.e. 50 mM Tris-HCl (pH 7.5), including 20 % glycerol, 0.5 µM FAD and FMN and 0.5 mM DTT. Following the buffer exchange, collected peaks were analysed by SDS-PAGE (Figure 3.18A). No protein could be detected so it seemed that all

protein was lost again during the buffer exchange to remove imidazole. Western blot, which is more sensitive than Coomassie staining, was needed to detect the purified CPRtr and indeed, degradation is observed (Figure 3.18B).

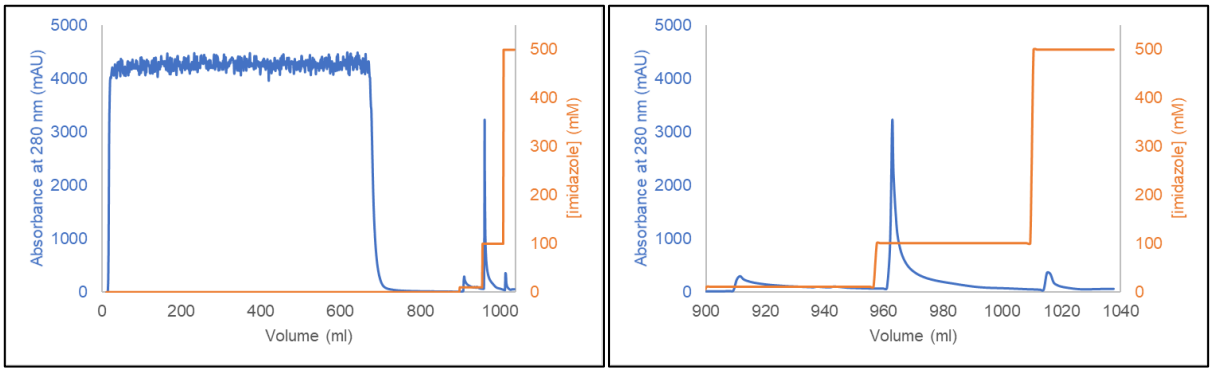

Figure 3.17: IMAC chromatogram of CPRtr purification from the medium after expression in the *HAC1* co-expressing strain in a 500 ml culture. 5 % glycerol was included in the IMAC buffers and pH adjustment in preparation of IMAC was performed more gently. The right chromatogram shows the elution profile in more detail.

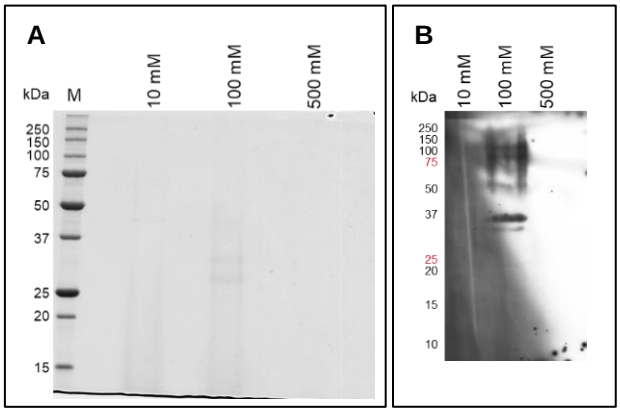

Figure 3.18: Analysis of IMAC elution peaks, desalted and concentrated after CPRtr (76 kDa) purification Figure 3.17. A: SDS-PAGE. B: Western blot.

Up until here, the buffer exchange was done using the Amicon® Ultra-15 centrifugal filter devices with a 10 kDa cutoff. Although we observed an intense peak eluting with 100 mM imidazole, after buffer exchange yield appeared to be very low. This indicates that the use of these centrifugal filter devices led to the protein loss. To this end, the same experiment was repeated, now performing ammonium sulphate precipitation of the collected elution peaks. The wild type strain producing CPRtr was taken along in parallel to check whether this adjusted IMAC protocol also led to an increased yield for this strain. Unfortunately, clogging of the sample pump filter again occurred when loading the collected medium after CPRtr production in the *HAC1* co-expressing strain on the IMAC column and the high yield could not be reproduced, based on the absorbance of the 100 mM elution peak (Figure 3.19). Even though no clogging occurred loading the medium of the wild type strain after CPRtr production, also here a decreased yield was observed (Figure 3.20), although not as significant as for the *HAC1* co-expressing strain. The HisTrap™ HP column was not stripped and recharged after the previous purification so this smaller decrease might be ascribed to a reduced capacity, even though the manufacturer states that the column can be used for five to seven purifications. As mentioned above, protein from the collected peaks were precipitated using ammonium sulphate

overnight. The next day, the obtained pellet was redissolved in the previously described storage buffer and analysed by SDS-PAGE. A sample of the collected peaks was taken before ammonium sulphate precipitation and analysed in parallel. The gel turned back empty after Coomassie staining. The protein yield thus was lower than the detection limit of Coomassie staining. A subsequent western blot analysis (Figure 3.21) showed that CPRtr indeed was recovered, be it in a very low yield. It was already observed that not one but two bands appeared when producing CPRtr in the wild type *P. pastoris* strain NRRL-Y-11430, which was confirmed here again. In the *HAC1* co-expressing strain, only one band appears. It thus seems that co-expression of *HAC1* results in correct and complete processing of the secretion signal, this in contrast to the wild type strain.

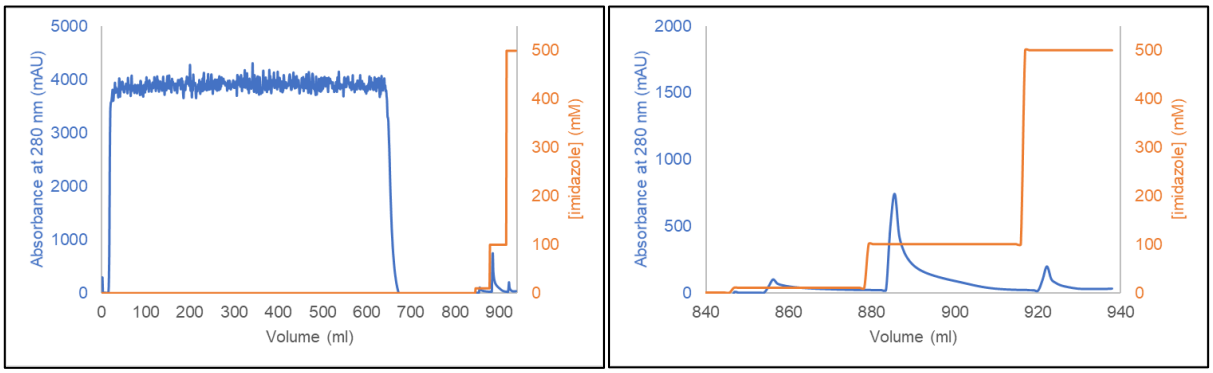

Figure 3.19: IMAC chromatogram of CPRtr purification from the medium after expression in the *HAC1* co-expressing strain in a 500 ml culture (replicate of Figure 3.17). The right chromatogram shows the elution profile in more detail.

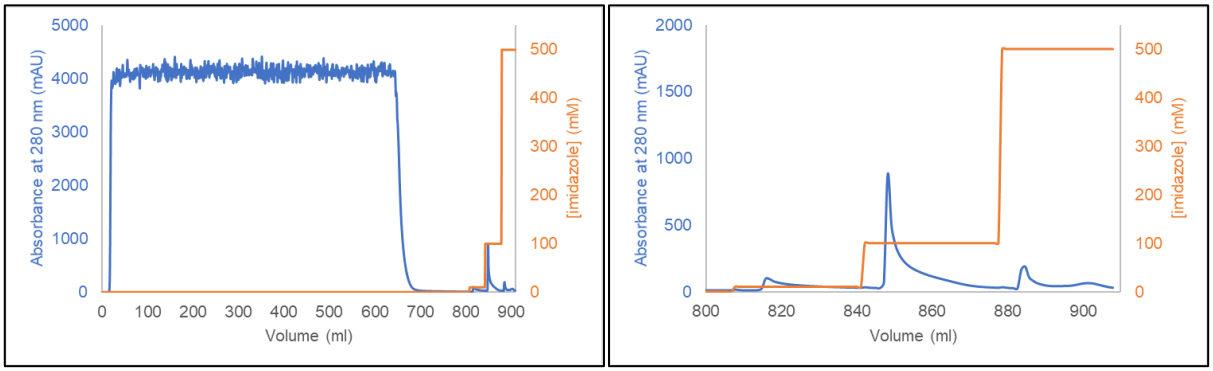

Figure 3.20: IMAC chromatogram of CPRtr purification from the medium after expression in the wild type strain in a 500 ml culture, using the adjusted IMAC protocol. The right chromatogram shows the elution profile in more detail.

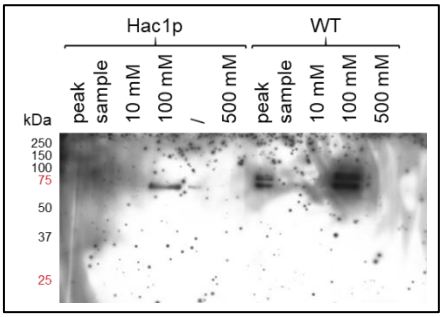

Figure 3.21: Western blot analysis of IMAC elution peaks, desalted and concentrated after CPRtr (76 kDa) purification Figure 3.19 and Figure 3.20.

CPRtr proved to be unstable and degrades or precipitates upon loading for IMAC purification or upon removal of imidazole, even in the presence of stabilizers like glycerol. Prior to any further efforts to purify the protein, we wanted to test whether it was active. Therefore, we precipitated the collected medium of both wild type and *HAC1* co-expressing *P. pastoris* producing CPRtr with ammonium sulphate, redissolved in the aforementioned storage buffer and desalted it. The protein samples were then analysed by western blot (Figure 3.22A). Only after CPRtr production in the wild type strain, the protein was successfully recovered. Using the nonphysiological redox partner cytochrome c, the obtained CPRtr proved to be active (Figure 3.22B). Due to the presence of the cofactors FAD, FMN and the additives glycerol and DTT, which interfere with protein quantitation methods like Bradford analysis, it was not possible to measure the protein concentration and thus specific activity. However, the activity per ml was calculated to be 0.18 U/ml. In a following step, the obtained CPRtr was added to microsomes containing CYPtr and the NADPH concentration was measured in presence and in absence of the substrate oleic acid (Figure 3.22C). No difference in NADPH consumption was observed between the two conditions. 100 µl CPRtr solution was added in this reaction, coinciding with 0.018 U and is probably not sufficient. For example, using the truncated CPR from *S. cerevisiae*, one unit is added in order to reconstitute CYP activity [316]. It could therefore not be concluded whether or not CYPtr was active and if CPRtr could reduce CYPtr.

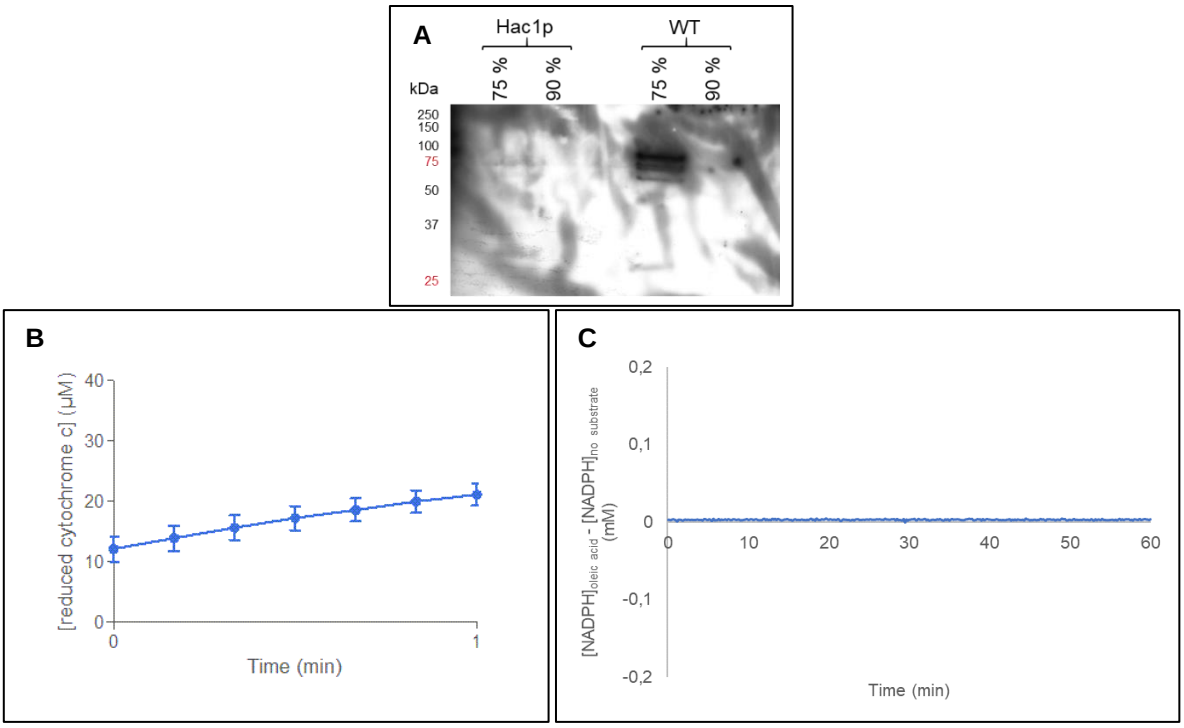

Figure 3.22: AS precipitation of protein from the medium after CPRtr (76 kDa) expression. A: Western blot analysis after desalting AS precipitated protein. B: Cytochrome c reductase activity assay using desalted AS precipitated protein from the wild type (WT) strain. C: NADPH oxidation assay using CYPtr-containing microsomes and including this crude CPRtr sample..

As the recombinant production of CPRtr was not the main goal of the project, further optimization was set aside and attention was given to the construction of a chimeric construct, together with the production of the nontruncated CYP52A13 in order to create a positive control and optimize from there.

3.3 Can fusion of the redox partner result in electron delivery to CYPtr?

The CYP with the highest catalytic activity reported to date, is the self-sufficient CYP102A1 from *B. megaterium*, and a high coupling efficiency is seen using this enzyme. Many artificial chimeric constructs have been created in order to mimic this self-sufficient CYP, some of which indeed showed improved catalytic properties [271]. Additionally, upon creating a self-sufficient CYP, only one enzyme is needed instead of two redox partners. Furthermore, it was hypothesized that fusion of a soluble reductase domain to the hydrophobic CYP might increase solubility, based on the successful solubilization of mammalian CYPs by truncation and fusion to BMR in the so-called molecular lego approach [273]. For these reasons, CYPtr and CPRtr were fused in order to create such a chimer. In accordance with CYP102A1, it was decided to fuse CPRtr with its N-terminus to the C-terminus of CYPtr, where the artificial N-terminus of CPRtr was removed and a glycine-rich linker was added. The *cprtr*-containing vector was linearized by PCR, substituting the NdeI restriction site by a NheI restriction site, enabling a double digest approach. Alternatively, *cyptr* and *cprtr* were fused by fusion PCR and subsequently ligated into an empty pPICZαB vector. Also CPEC was attempted. Both the substitution and fusion PCR approach led to successful cloning, whereas CPEC did not.

The obtained plasmids were linearized and transformed in *P. pastoris* NRRL-Y-11430. A small-scale expression test was performed using three colonies obtained from each molecular cloning approach. Medium, lysate and microsome samples were collected and analysed by western blot (Figure 3.23A, B and C, respectively). The chimeric construct showed to be expressed well since a highly intense band is observed in the higher MW region of the blot (expected MW of 131 kDa, or 141 kDa, when the α-factor secretion signal is still included). The protein again stays bound to the microsomal membrane and no secretion to the medium is observed. Next to the intense band in the region of 131 kDa, additional bands are seen, indicating degradation to a large extent in the lysate samples (Figure 3.23B). In the purified microsomal samples (Figure 3.23C), more discrete bands are present next to the one in the high MW region. For instance, a discrete band is observed between 50 kDa and 75 kDa, coinciding with the MW of CYPtr. The glycine-rich linker seems thus prone to proteolysis. Since the N-terminal anchor is removed, it was wondered whether the chimeric product was peripherally bound to the ER or still integrally bound. To answer this question, the microsomes were washed with a high salt buffer and the supernatant of this wash was analysed by western blot (Figure 3.23D). A clean band around the expected MW of the chimeric construct was present. Thus, even though the truncation did not lead to complete solubilization, the enzyme is only peripherally bound and could potentially be purified using high salt buffers, in absence of detergent. We concluded that the chimeric product is successfully produced at some extent.

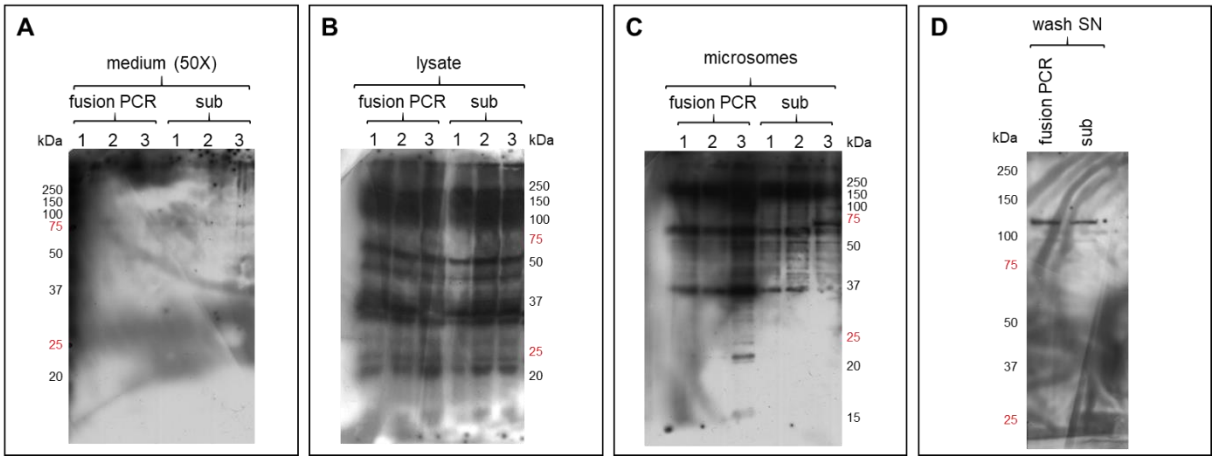

Figure 3.23: Western blot after small-scale expression test of three colonies picked after *cyptrcprtr* transformation in *P. pastoris*. A: concentrated medium samples. B: cell lysate. C: isolated microsomes. D: supernatant of microsome wash with high salt buffer. The fusion enzyme CYPtrCPRtr (131 kDa, 141 kDa in case of an unprocessed signal sequence) is produced but not secreted to the medium. The fusion enzyme is retained in the microsomal membrane. Solubilization was accomplished to some extent, as the fusion enzyme was retrieved in the soluble supernatant after a wash step with a high salt buffer.

In a similar fashion as described for the CYPtr-containing microsomes, an activity assay was performed using microsomes isolated from a 100 ml culture to verify activity of the chimeric construct. Oleic acid conversion to 1,18-octadecenedioic acid was checked by GC-MS analysis. Unfortunately, no DCA was found in the GC chromatogram (data not shown). Derivatization by methylation prior to GC-MS is only useful to show FAs and DCAs. No OHFAs could be detected in this way. Although Eschenfeldt et al report that CYP52A13 is able to overoxidize oleic acid to the corresponding DCA, not all oleic acid is converted to the corresponding DCA and also 18-hydroxyoleic acid is formed [24]. Therefore, the GC-MS protocol was adjusted and now silylation instead of methylation was performed. Additionally, an NADPH regeneration system was included in the enzymatic reaction. After the conversion assay with microsomes containing CYPtrCPRtr, an additional peak was observed at a retention time of 9.42 min (Figure 3.24A), the region where OHFAs are expected based on standards 16-hydroxypalmitic acid (RT = 8.9 min) and 12-hydroxystearic acid (RT = 9.3 min). Unfortunately, this did not correspond to a OHFA. The spectrum was searched against the NIST database, which identified the compound as myo-inositol, 1,3,4,5,6-pentakis-O-(trimethylsilyl)-, bis(trimethylsilyl) phosphate.

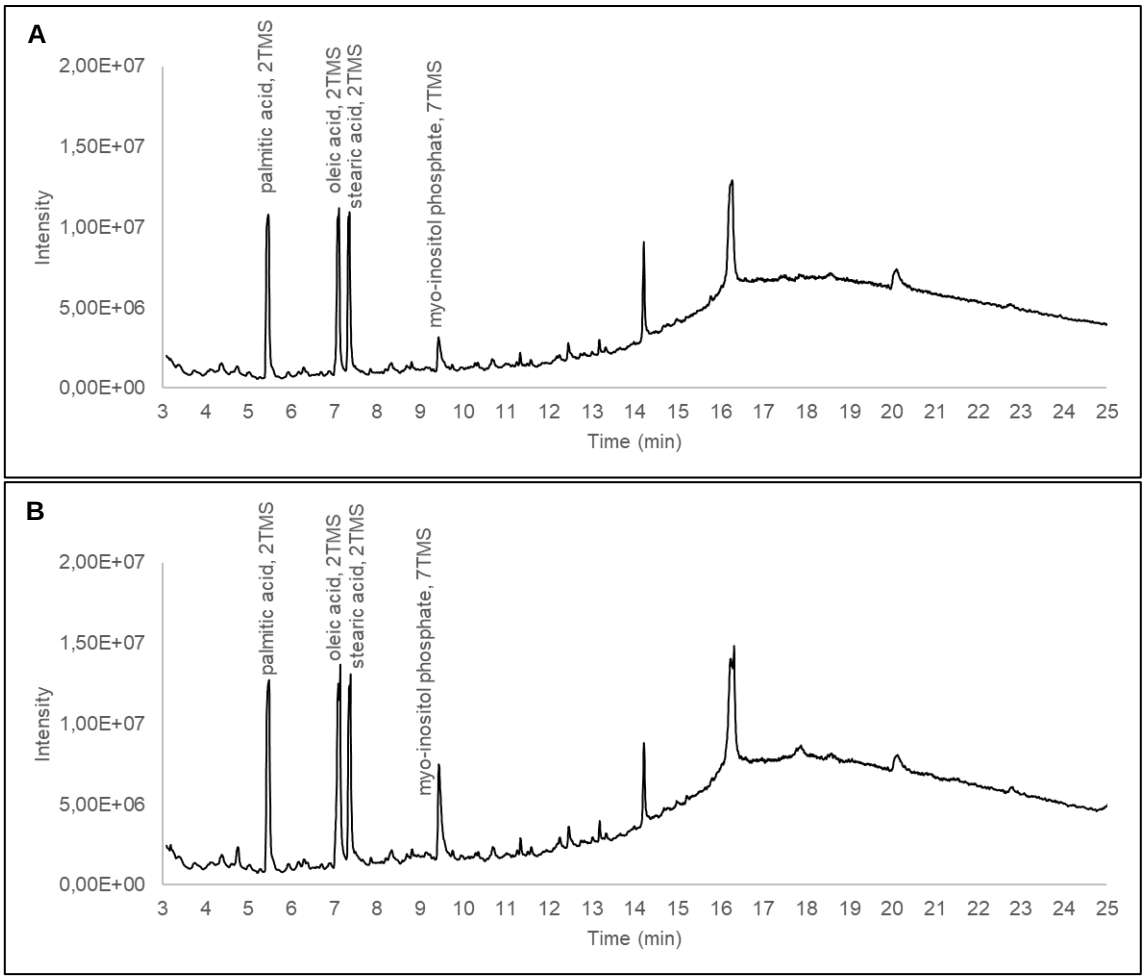

Figure 3.24: GC chromatogram of silylated extracts after microsome activity assay with chimeric construct CYPtrCPRtr. Only peaks identified as silylated compounds, based on the presence in the MS spectrum of the characteristic peaks at m/z 73 and 75, are annotated. No conversion of oleic acid to the corresponding OHFA (RT unknown, no standard available) or DCA (expected RT = 11.2 min) was observed. A: chimer. B: negative control.

Next to these oleic acid conversion assays, followed by GC-MS analysis of the diethyl ether extracted products, NADPH oxidation was measured spectrophotometrically in the same way as described for CYPtr. In another setup, the difference in NADPH consumption in presence and in absence of the substrate oleic acid was measured. In both cases, no increased NADPH oxidation was observed (data not shown). This coincides with the fact that no conversion of oleic acid occurred. Furthermore, no substrate-induced uncoupling through one of three possible uncoupling reactions seems to be happening.

It was tested whether the reductase domain was active to see if the lack of activity was not due to an inactive redox partner. To this end, the nonphysiological redox partner cytochrome c was used. Microsomes were isolated from a 100 ml culture after CYPtrCPRtr production. In parallel, microsomes were isolated from the wild type NRRL-Y-11430 strain and from a culture after CYPtr production. In all cytochrome c reductase assays, the same amount of microsomal protein, i.e. 0.1 mg, was added and it was investigated whether the microsomes collected after CYPtrCPRtr production showed an increased reductase activity (Figure 3.25). Indeed the microsomes with the chimer showed an increased reductase activity compared to both the wild type microsomes and the microsomes with CYPtr. Within seconds, equilibrium is reached in

case of the chimer. The reductase domain of the chimer is thus active. Of note, recombinantly producing CYPtr in *P. pastoris* seems to minorly influence the reductase activity of the microsomal fraction. Performing an unpaired t-test yields a p-value of 0.0325, showing that the slope is significantly more steep in case of CYPtr-containing microsomes, compared to the wild type microsomes.

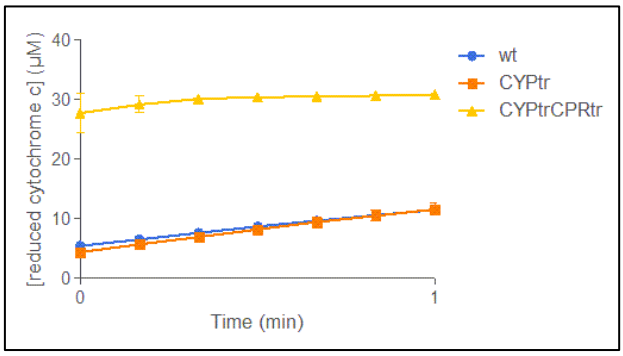

Figure 3.25: Cytochrome c reductase assay of microsome samples after CYPtrCPRtr production compared to CYPtr production and no recombinant production (i.e. wt). CYPtrCPRtr-containing microsomes reach equilibrium within seconds, which is not the case for the two other microsome samples. The average of three technical replicates is shown.

Many examples are reported where the reductase domain of CYP102A1 from *B. megaterium* is used in an artificial fusion enzyme [271]. In our effort to produce a self-sufficient DCA producing enzyme, this reductase domain (BMR) was used to create an alternative chimer. BMR was N-terminally fused to the C-terminus of CYPtr using the natural linker of CYP102A1, consisting of residue 459 to 478 according to Sevrioukova et al [190]. CPEC was applied for molecular cloning of *bmr* into the *cyptr*-containing pPICZαB vector. Again, a small scale expression test with three picked *Pichia* colonies was performed, followed by a western blot analysis of the medium, lysate and microsomal fraction (Figure 3.26A, B and C, respectively). Similarly as for the CYPtrCPRtr chimer, the enzyme remained bound to the microsomal membrane and degradation was observed.

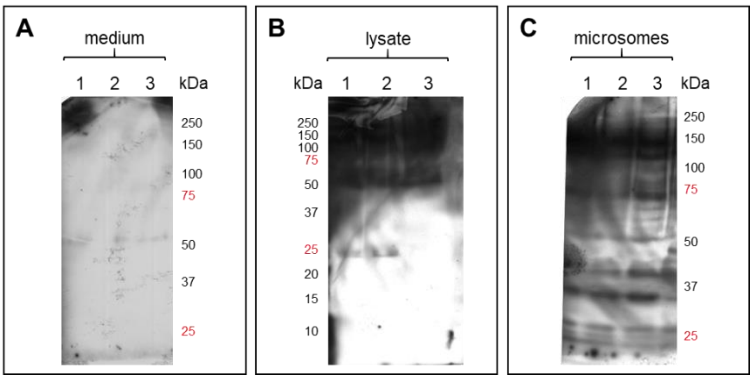

Figure 3.26: Western blot after small-scale expression test of three colonies picked after *cyptrbmr* transformation in *P. pastoris*. A: concentrated medium samples. B: cell lysate. C: isolated microsomes. The fusion enzyme CYPtrBMR (123 kDa, 132 kDa in case of an unprocessed signal sequence) is produced but not secreted to the medium. The fusion enzyme is retained in the microsomal membrane.

The microsomes were checked for activity in the same way as described for CYPtrCPRtr. Product was extracted with diethyl ether after an enzymatic assay for subsequent methylation and GC-MS analysis, and an NADPH oxidation assay was performed as well (data not shown).

Both assays showed that no oleic acid was converted nor substrate induced uncoupling occurred. Presence of the CYPtrBMR fusion enzyme in the microsomes did lead to increased reductase activity using the nonphysiological redox partner cytochrome c (Figure 3.27).

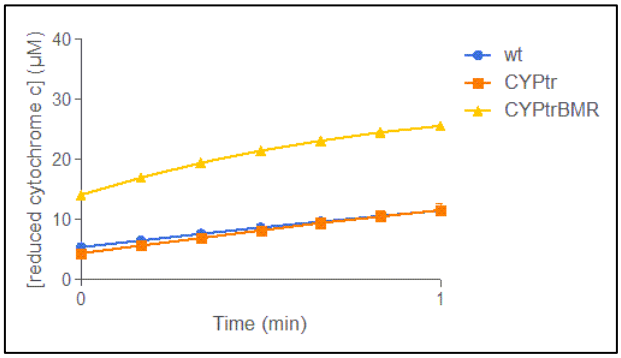

Figure 3.27: Cytochrome c reductase assay of microsome samples after CYPtrBMR production compared to CYPtr production and no recombinant production (i.e. wt). CYPtrBMR-containing microsomes showed a higher reductase activity. The average of three technical replicates is shown.

3.4 Are the truncations and/or presence of the α-factor secretion signal leading to inactivity?

3.4.1 Recombinant production of the nontruncated CYP52A13

The first exploratory experiments using CYPtr showed that the truncated enzyme remained bound to the microsomes and was not secreted to the medium. In the enzymatic assays, no conversion to the desired DCA occurred, nor was there any increased substrate-induced NADPH consumption. In the chimeric constructs, we observed CPR activity, and thus it was concluded that lack of activity is due to inactive CYPtr. It was also attempted to recombinantly produce CYP52A13 in a nontruncated form (CYPnt) in *P. pastoris*. In this way, it was investigated whether or not the truncations led to the loss of activity. Using a nontruncated form would additionally allow for the optimization of a self-sufficient chimer in its natural environment of the ER membrane, before continuing with solubilization efforts. Furthermore, a positive control could be created in this way.

CYPnt_pPICZA was linearized and transformed in *P. pastoris* NRRL-Y-11430. Six colonies were picked and tested in a small scale expression test. No Histag is present so not western blot but Coomassie staining of the SDS-PAGE gel was performed (Figure 3.28). As there are no truncations, CYPnt should reside in the microsomal fraction. To identify the gel band, coinciding with CYPnt, a microsome sample of the empty wild type *P. pastoris* strain was taken along. Comparing the profile of the six different colonies to the wild type microsome sample shows a different pattern between 50 and 75 kDa, the region where CYPnt (60 kDa) is expected. Gel bands were cut out in this region and analysed by MALDI-TOF MS (data not shown). The more intense band around 75 kDa was identified as the peroxisomal alcohol oxidase 1 (AOX1). Upon methanol induction, AOX1 can account for up to 30 % of the total soluble protein [312]. It is thus not surprising that AOX1 is contaminating isolated fractions. The two bands around 50 kDa were identified as the mitochondrial ATP synthase subunits α and β (ATP1 and ATP2). The ATP synthase subunits are found in the inner membrane of the mitochondria and were

reported to contaminate isolated microsomal fractions before [325]. The intermediate bands could not be identified. However, if CYPnt would be produced to a high yield, this should be visible as a prominent band in the 50-75 kDa region on gel. Efforts were done to improve the microsome isolation protocol in order to further enrich the microsomal proteins leading to the identification of recombinant CYPnt. Including additional wash steps of the microsomal fraction resulted in identification of several additional proteins in the aforementioned MW region, however, not leading to the identification of CYPnt (data not shown). Moreover, enzymatic cell lysis instead of mechanical cell lysis with glass beads was performed in order to leave peroxisomes and mitochondria intact and isolate these organelles using intermediate centrifugation speeds. Alternatively, a prolonged expression phase of 72 h and 96 h was executed in an attempt to increase the CYPnt yield. However, CYPnt could not be identified in any case (see Addendum).

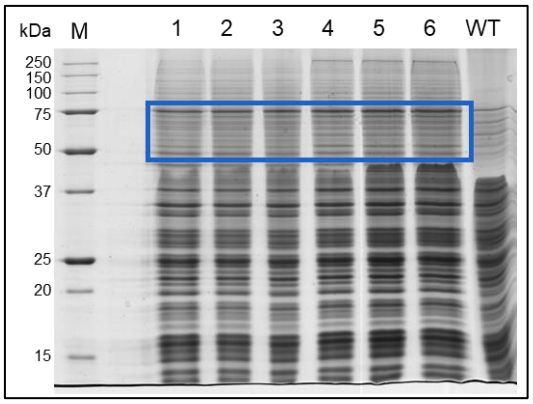

Figure 3.28: SDS-PAGE analysis of 20 µg microsomal protein after small-scale expression test of six colonies picked after *cprnt* transformation in *P. pastoris*. Gel bands in the blue rectangle were cut out for MALDI-TOF MS. CYPnt (60 kDa) was not identified in the analysed gel bands.

In parallel to the MALDI-TOF MS identification efforts, enzymatic assays were performed as an alternative way to prove production of an active CYPnt. At first, it was again assumed that the endogenous CPR from *P. pastoris* could serve as the alternative redox partner and it was tested whether whole cells could be used to assess activity instead of performing *in vitro* assays using isolated microsomes. On the one hand, the substrate oleic acid was added during the expression phase. On the other hand, cells were collected after the expression phase, washed and resuspended in phosphate buffer for a subsequent whole-cell bioconversion assay. After bioconversion, the products were extracted with diethyl ether and methylated for subsequent GC-MS analysis. Oleic acid was completely consumed after 24 h and no corresponding DCA appeared in the GC chromatogram (Figure 3.29). The same was observed in case of the *in vitro* whole-cell bioconversion assay, i.e. complete consumption of oleic acid without the appearance of DCA (data not shown). Although also 18-hydroxyoleic acid could be the final product, which is not detected with the applied GC-MS protocol, it is more likely that the oleic acid was completely consumed by β-oxidation in the peroxisomes. Upon methanol induction, it is known that the peroxisomes proliferate. Also FAs can be used as the sole carbon source and addition of FAs as well leads to peroxisome proliferation [326]. Therefore, further enzymatic assays were performed with isolated microsomes.

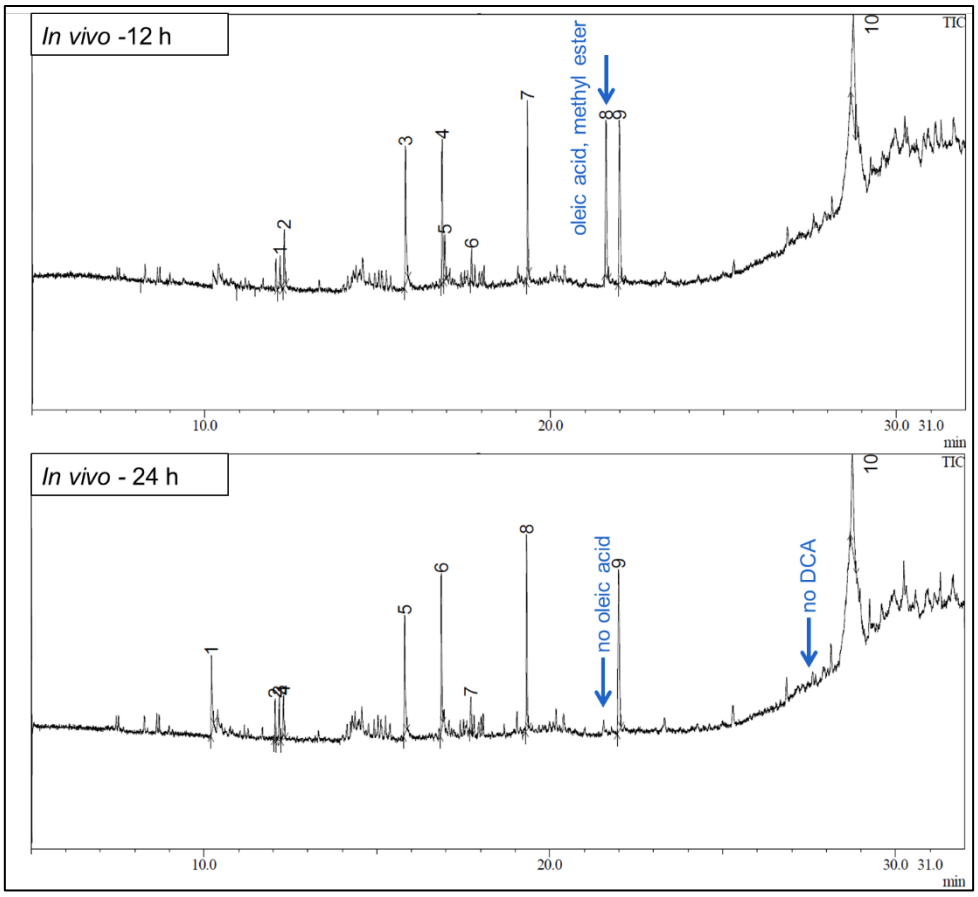

Figure 3.29: GC chromatogram of samples taken during an *in vivo* bioconversion assay. Oleic acid was added to the *P. pastoris* culture during the *cypnt* expression phase and samples were taken every 12 h. Oleic acid was completely consumed after 24 h without DCA formation (expected RT = 27.45 min).

Microsomes were isolated and used in different enzymatic assays. First NADPH oxidation was monitored in presence and in absence of the substrate oleic acid. No difference in the rate of consumption was observed (data not shown). In another assay, the diethyl ether extract, collected after conversion, was silylated and analysed by GC-MS. No 18-hydroxyoleic acid, nor 1,18-octadecendioic acid was found in the chromatogram (data not shown). Of note, in both assays, recombinant CPRtr was added. As described before, the isolated CPRtr had a low activity per ml so it can be disputed whether this was enough to deliver the required electrons to CYPnt.

Using SDS-PAGE and subsequent MALDI-TOF MS of excised gel bands did not result in the identification of CYPnt. Also the performed enzymatic assays came back negative. Therefore, it was decided to add a Histag C-terminally to CYPnt enabling the use of the more sensitive western blot analysis. Additionally, it was asked whether the expression yield could be improved by altering the translation initiation sequence. Invitrogen proposes the following sequence as a suitable translation initiation sequence for high expression (G/A)NNATGG, that is the so-called Kozak sequence. First, the Histag-containing enzyme was constructed by PCR using back-to-back primers, followed by NheI digest and ligation. Back-to-back primers were then applied as well in order to modify the translation initiation sequence, resulting in the following sequence GAGATGGGT. A Gly residue was inserted between Met and Thr with the aim of introducing the guanine base following the ATG start codon. Three colonies were picked

upon transformation of *P. pastoris* with the linearized plasmid in each case for a small scale expression test. Microsomes were subsequently collected and analysed by western blot (Figure 3.30). Only a very light signal is observed in the 50-75 kDa region in case of CYPntHis. This coincides with the expected MW of 60 kDa. This very low signal shows that the expression yield is low and explains why CYPnt proved so difficult to find on SDS-PAGE gel. After introduction of the Kozak sequence, no enzyme was present in the microsomal fraction, based on western blot analysis. It could thus be concluded that the introduction of this Kozak sequence did not improve the protein yield.

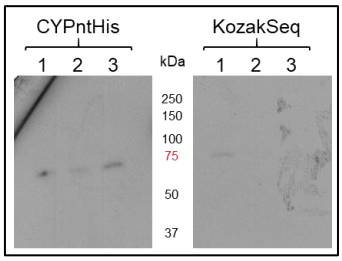

Figure 3.30: Western blot after small-scale expression test of three colonies picked after both *cypntHis* and *cypntHis_KozakSeq* transformation in *P. pastoris*. CYPntHis (60 kDa) was only produced in a low amount. No CYPntHis is present after altering the translation initiation sequence (KozakSeq).

Western blot analysis resulted only in a very low signal so other means of CYPnt identification and possibly quantification were considered. In this regard, an LC-MS method, multiple reaction monitoring (MRM), was applied for CYPnt detection and quantification in order to investigate which constructs lead to the highest enzyme yield. For MRM, there is no need for a Histag so the original CYPnt enzyme was again used. As the Kozak sequence did not lead to an increased CYPnt yield, another translation initiation sequence was investigated. The nontruncated form of CPR-a (from here on now referred to as CPRnt) shows to be produced well in *P. pastoris* (discussed below) and it was therefore chosen to substitute the 5' region in *cypnt* with the 5' region from *cprnt*, that is ACCATGGCA. Of note, this sequence also coincides with the suggested Kozak sequence as described above. This sequence was introduced using back-to-back primers in the same way as performed for the previous Kozak sequence. After transformation of the linearized plasmid in *P. pastoris*, a 100 ml expression test was performed, taking along the original CYPnt construct. Microsomes were isolated and prepared for MRM. In the sample preparation, BSA was added as an internal standard and the AUC was normalized using this internal standard. The AUC of CYPnt_TrlInCPRnt was adjusted relative to CYPnt and the results are shown in Figure 3.31. From this it was concluded that again, modification of the translation initiation sequence led to a decreased instead of an increased yield.

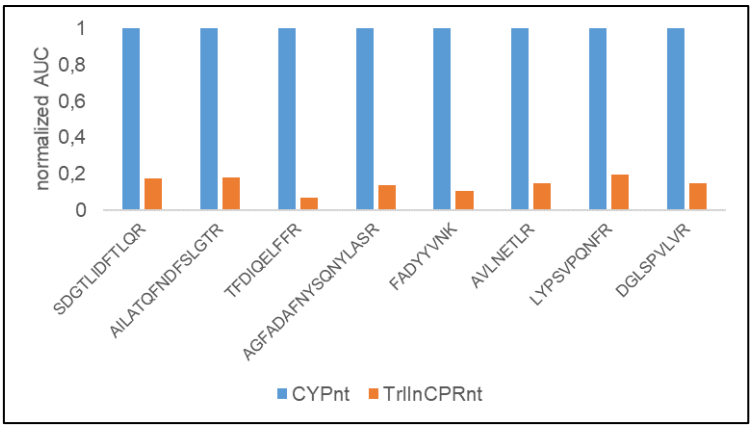

Figure 3.31: Relative abundance of CYPnt before and after introducing the translation initiation sequence of CPRnt, determined by MRM. The total AUC for each peptide was calculated and normalized to BSA. The normalized AUC for each CYPnt peptide taken along in de analysis is shown. The normalized AUC of CYPnt was corrected to 1 and the construct with the alternative translation initiation sequence was corrected relative to CYPnt. The alternative translation initiation sequence leads to a decreased instead of an increased yield.

3.4.2 Providing the redox partner for the nontruncated CYP52A13

3.4.2.1 Chimeric construct

CYP needs a redox partner for the delivery of electrons and for initial enzymatic assays, it was hypothesized that the endogenous CPR of *P. pastoris* might serve to this purpose. As no active system was obtained for designs based on this assumption, recombinant production of CPRtr was explored for the external addition in enzymatic assays. As described above, only low activity samples could be obtained and addition to either CYPtr or CYPnt, present in the microsome, did not lead to oleic acid-converting microsomes. Therefore, an artificial chimer was created between the CYPnt and BMR in the same way as described for CYPtr. Expression tests were performed in parallel with CYPnt and expression was difficult to prove for the chimeric construct as well. It was hypothesized that also in this case, only a very low yield was obtained. This was evidenced by a cytochrome c reductase assay, shown in Figure 3.32. A 100 ml expression test was executed after which the microsomes were isolated, taking along CYPnt. The reductase activity of 0.1 mg microsomal protein after CYPntBMR production was compared to the reductase activity of 0.1 mg microsomal protein after CYPnt production and showed only a minor increase. Referring to Figure 3.27, presence of the BMR domain results in a marked reductase activity increase in the CYPtrBMR chimeric construct. A similar increase in reductase activity is not observed here, thereby confirming the low enzyme yield. Due to this low yield, no further efforts were done to construct CYPntCPRtr chimers.

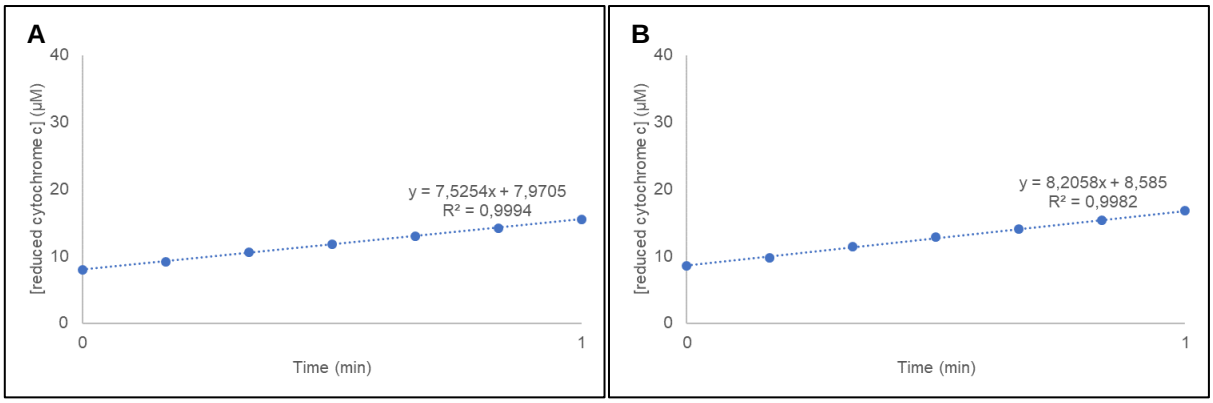

Figure 3.32: Cytochrome c reductase assay of microsome samples. A: CYPnt. B: CYPntBMR.

3.4.2.2 Co-expression of the nontruncated CPR-a

No active enzyme was yet obtained and several reasons could be ascribed to that. No suitable redox partner was present, that is neither the endogenous CPR from *P. pastoris* nor CPRtr or the BMR domain were sufficient. The truncations might lead to loss of activity, which could not be verified using the nontruncated form due to its low production yield. An attempt was done to prove the activity of CYP52A13 in *P. pastoris* by creating a positive control where CYPnt was co-expressed with a nontruncated form of CPR-a from *C. viswanathii* (CPRnt) in the ER of *P. pastoris*. It was hypothesized that co-expression of the natural redox partner CPRnt might be beneficial for CYPnt production as it has been reported that CYP production and stability are dependent on the co-expressed CPR [237], [327]. Interestingly, Gudiminchi et al optimized a high-throughput screening method for CYP production in *P. pastoris* by measuring the CO-bound difference spectrum of whole cells. They applied their protocol to investigate in which case CYP52A13 yields were the highest and saw that co-expressing CPR increased the CYP yield [327]. To this end, a co-expression plasmid was created where the *cprnt* expression cassette was cloned into the *cypnt* vector using a Golden Gate-inspired molecular cloning approach. To be able to linearize the obtained plasmid with the PmeI restriction enzyme for efficient integration into the genome of *P. pastoris*, the PmeI restriction site was removed from the *cprnt* vector by site directed mutagenesis before cloning the *cprnt* expression cassette into the *cypnt* vector. For rapid detection, constructs including a C-terminal Histag were used. A 100 ml expression test was performed, taking along the CYPntHis construct. Microsomes were isolated and it was already evident from the collected pellet that CPRnt was present as the colour was markedly different from the microsomes containing CYPntHis alone (Figure 3.33A). The pellet appears to be more green-brown, the greenish colour coinciding with the presence of a blue semiquinone flavin and a yellow fully oxidized flavin [183]. Western blot analysis (Figure 3.33B) shows a very intense signal, not only at the expected MW of 78 kDa (CPRnt) and 60 kDa (CYPnt), but a smear is seen. Therefore, it could not be concluded whether one of the two or both enzymes were produced and what the ratio between the two enzymes was.

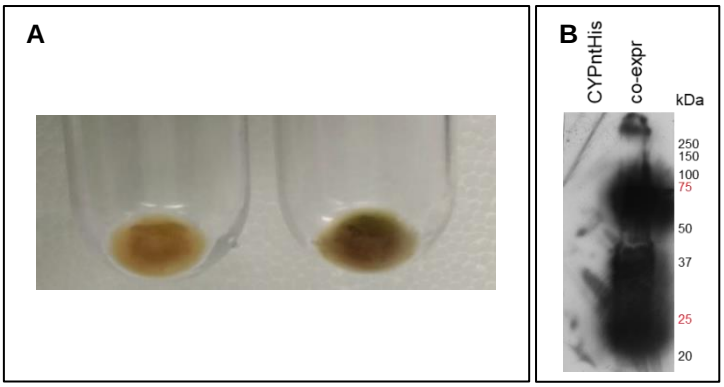

Figure 3.33: A: Microsomal pellet isolated after (left) CYPntHis production and (right) CYPntHis-CPRntHis production. B: Western blot analysis of these isolated microsomes. No signal is seen for CYPntHis (60 kDa). After co-expression, a smear is observed, instead of two distinct signals coinciding with CYPntHis (60 kDa) and CPRntHis (78 kDa).

The next step was to use these microsomes in an *in vitro* enzymatic assay. Not one but five different enzymatic assays were set up, selected from different sources, all assaying ω-hydroxylation (see 2.13.1). Subsequent to each assay, the extracted products were analysed by GC-MS upon silylation. Unfortunately, not one chromatogram showed the presence of 18-hydroxyoleic acid or 1,18-octadecenedioic acid (data not shown). Here again it could be questioned whether this was actually due to an inactive CYPnt. The western blot in Figure 3.33 shows an intense signal together with a lot of degradation whereas CYPntHis shows no signal at all. Even though CPRntHis co-expression could lead to an increased yield of CYPntHis, it was considered that most of the signal came from CPRntHis, based on other co-expression experiments (see below). This would mean that instead of a CYPntHis:CPRntHis ratio higher than one, as present in natural systems, the ratio is lower than one and an excess of CPRntHis is present. Indeed, Paddon et al reported it to be advantageous to express *cpr* from a weak promotor [132].

3.4.3 Evaluation of CYP52A13 constructs by mass spectrometry

Throughout, many different constructs were made, which were characterized by a set of different techniques and assays. However, none of these tests allows to compare the actual production levels of recombinant CYP in the different set ups. Therefore, we designed a mass spectrometric assay implementing MRM to track the presence of CYP-derived peptides in tryptic digestions from protein extracts from the producing strains. Therefore, a milli-scale experiment was set up, where 2 ml cultures were grown in a 24-well deep-well plate. This allowed for the analysis of all constructs in parallel and in triplicate. Microsomes from the *P. pastoris* producing CYPtr were included in this experiment as well. Subsequent to growth and recombinant protein production, microsomes were collected and prepared for quantification by MRM. The abundance of eight peptides was quantified by determining the AUC and normalization to BSA (Figure 3.34). For each peptide, the abundances of the different constructs were significantly different from one another, as determined by one-way ANOVA ($p < 0.0001$ for all eight peptides). Tukey's multiple comparison test further specified that the CYPtr construct was significantly more abundant than all CYPnt constructs, and this for every peptide. No significant differences were seen between the different CYPnt constructs, based on Tukey's multiple comparison for the peptides shown in Figure 3.34A, B, C, D, E and H. In case of the peptide LYPSVPQNFR, the peptide abundance for CYPnt was significantly higher than for the CYPntBMR chimer, the CPRnt co-expression construct and both CYPnt constructs with an alternative translation initiation sequence (KozakSeq and TrlInCPRnt) (Figure 3.34F). In case of the peptide DGLSPVLVR, CYPnt was significantly more abundant than the CPRnt co-expression construct (Figure 3.34G). Although other differences were not significant, it could be seen that the applied CYPnt modifications did not lead to an increased, but rather to a decreased yield. It must be noted that peak tailing was observed in the MRM results, probably due to the many contaminants in the microsomal samples, e.g. presence of membrane fractions. However, the same general trend was observed for all eight peptides taken along in the analysis. Based on this evaluation, it could be concluded that the truncation and/or addition of the N-terminal tag leads to an increased protein yield. Although it must be considered that the truncation might enhance solubility and extraction efficiency in the sample preparation leading up to MRM, this increased yield was suspected based on earlier experiments. The western blot shows an intense signal for CYPtr in the microsome and in case of the CYPtr chimers, clear reductase activity was observed, which was not the case for the CYPnt chimer.

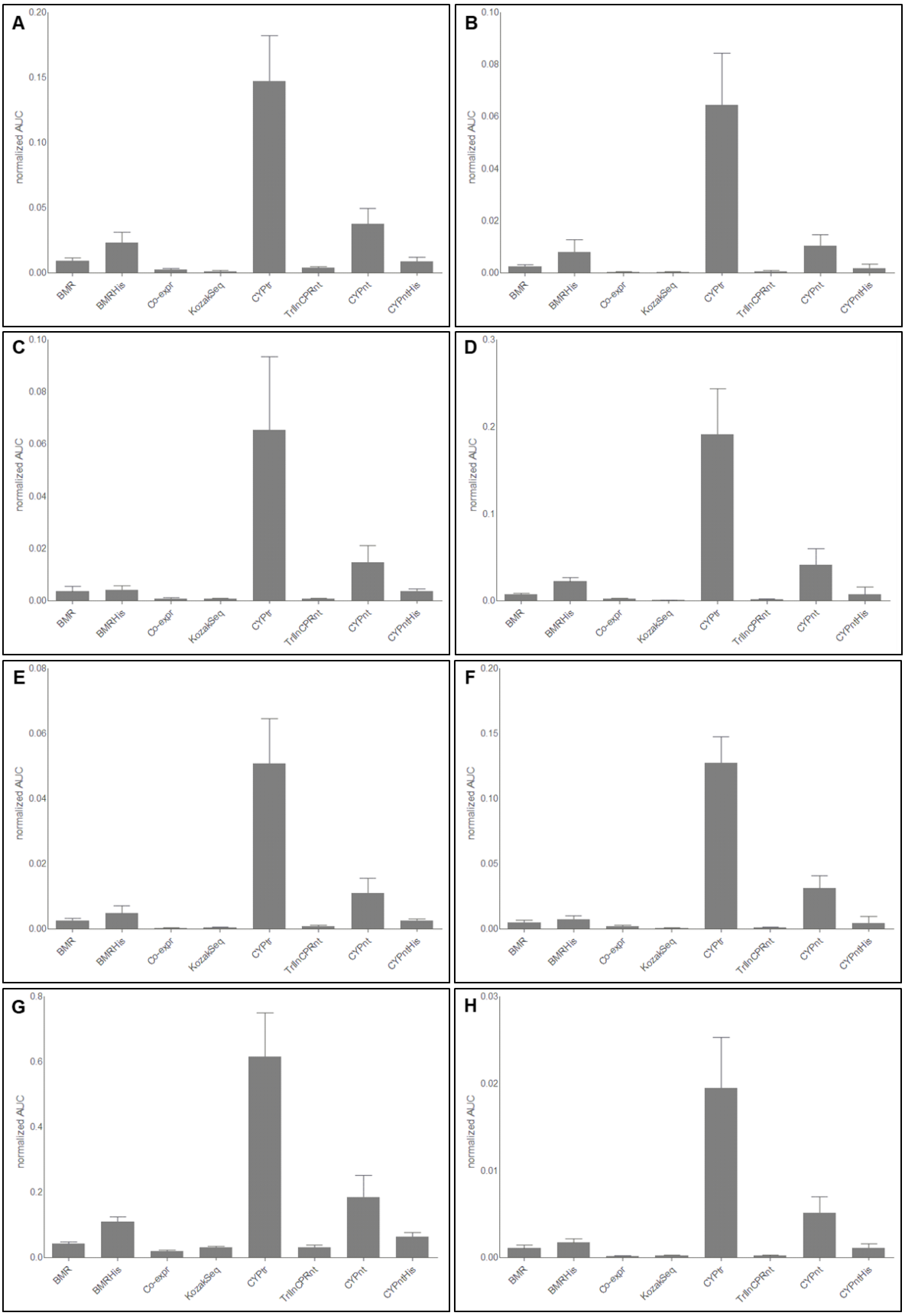

Figure 3.34: Abundance of different CYP constructs in de microsome, determined by MRM. The total AUC for each peptide was calculated and normalized to BSA. The normalized AUC (average of three biological replicates) for each CYP peptide taken along in de analysis is shown. A: SDGTLIDFTLQR, B:AILATQFNDFSLGTR, C: TFDIQELFFR, D: FADYYVNK, E: QDGYVFLYELVK, F: LYPSVPQNFR, G: DGLSPVLVR, H: LGWAFLPFNGGPR.

3.4.4 Recombinant production of CYPtr in a *HAC1* co-expressing strain

The use of the *HAC1* co-expressing strain showed promise in recombinant CPRtr production in that a markedly higher yield could be obtained, based on the peak area of the 100 mM imidazole fraction containing CPRtr. Additionally, instead of two bands, only one band appeared in both SDS-PAGE and western blot analysis, indicating that the α-factor secretion signal was correctly processed only in this *HAC1* co-expressing strain. After producing CYPtr in the wild type strain NRRL-Y-11430, the enzyme was not secreted and appeared at a higher MW, indicating that the secretion signal was not processed. Therefore, *cyptr* was transformed in this *HAC1* co-expressing strain to see if also for this enzyme, the processing of the secretion signal was aided. Three colonies were taken along in a small scale expression test. A medium sample was concentrated and analysed by western blot, together with the lysate and the isolated microsomes (Figure 3.35). Remarkably, CYPtr was found in the medium in case of the second colony. A band is seen around the expected MW of 58 kDa, whereas this is not visible in case of the first and third colony. However, the enzyme is still partially retained in the microsomal fraction.

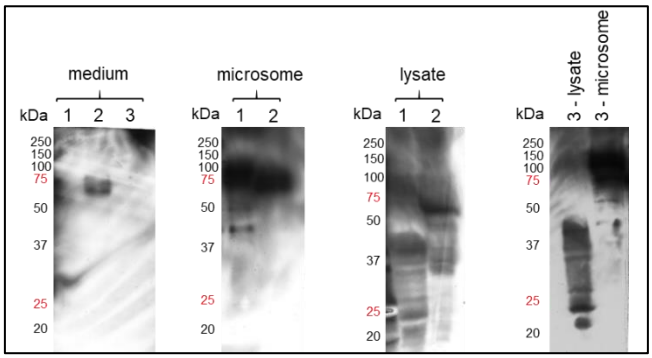

Figure 3.35: Western blot of concentrated medium, lysate and microsome samples after small-scale expression test of three colonies picked after *cyptr* transformation in the *HAC1* co-expressing strain. CYPtr (58 kDa, 67 kDa in case of an unprocessed signal sequence) is successfully produced and secreted in case of colony 2. In colony 1 and 3, no secreted enzyme is observed and CYPtr is largely degraded.

This observation prompted us to evaluate whether the chimeric constructs from 3.3 could be secreted as well using the *HAC1* co-expressing strain. Three colonies of each chimeric construct obtained after transforming this *P. pastoris* strain were picked for a small-scale expression test. A medium sample was concentrated and microsomes were isolated for western blot analysis (Figure 3.36). Unfortunately, no enzyme was observed in the medium and the chimers were still retained in the microsomes.

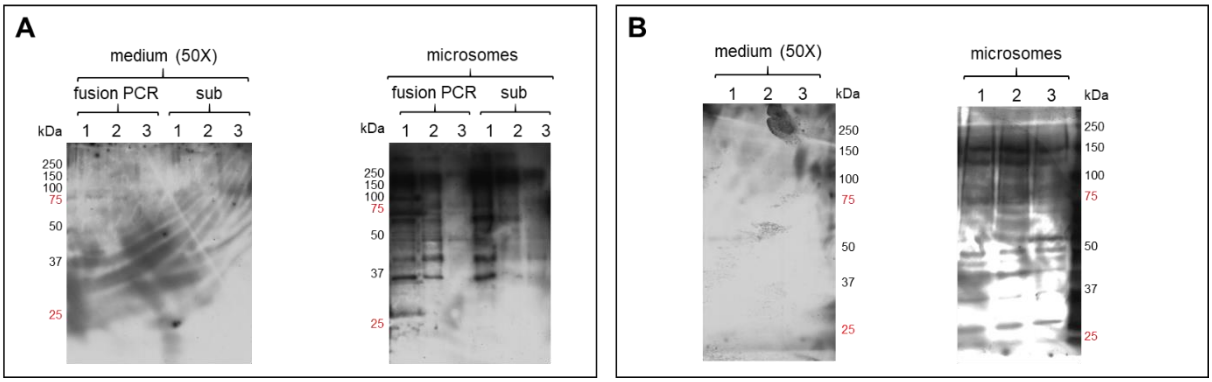

Figure 3.36: Western blot of concentrated medium and microsome samples after small-scale expression test of three colonies picked after chimer transformation in *HAC1* co-expressing *P. pastoris*. A: CYPtrCPRtr (131 kDa, 141 kDa in case of an unprocessed signal sequence), obtained by fusion PCR and by the substitution approach. B: CYPtrBMR (123 kDa, 132 kDa in case of an unprocessed signal sequence). Both chimeric constructs were not secreted but were retained in the microsomal fraction.

We continued investigating colony 2, able to secrete CYPtr into the medium. Cells were grown in a 500 ml culture, followed by IMAC purification of the medium (Figure 3.37) and the different elution peaks were analysed by SDS-PAGE (Figure 3.38). No clear band was observed around the MW of 58 kDa, the expected MW of secreted CYPtr. In both the 10 mM and 100 mM imidazole elution peak, an intense band was observed at a lower MW. In the 100 mM imidazole elution peak, two additional bands were seen. All these gel bands were cut out for MALDI-TOF MS. Unfortunately, CYPtr could not be identified and elution peaks contained contaminants only.

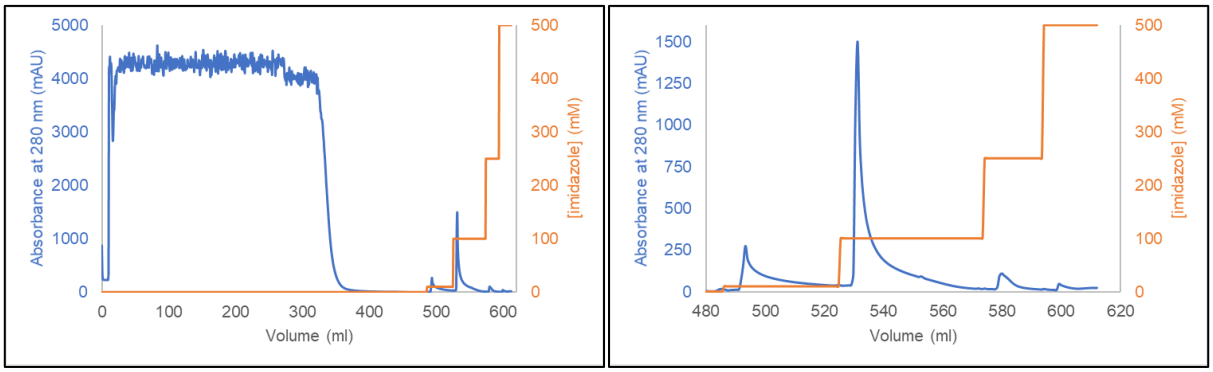

Figure 3.37: IMAC chromatogram of CYPtr purification from the medium after expression in a 500 ml culture. The right chromatogram shows the elution profile in more detail.

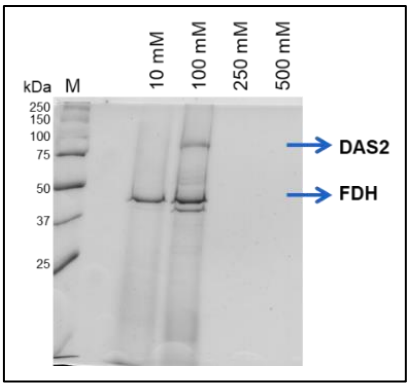

Figure 3.38: Analysis of IMAC elution peaks, desalted and concentrated after CYPtr (58 kDa) purification Figure 3.37. The gel bands were cut out and analysed by MALDI-TOF MS. The higher MW gel band in the 100 mM imidazole elution peak was identified as dihydroxyacetone synthase variant 2 (DAS2).The gel band in between 37 kDa and 50 kDa, present in both the 10 mM and 100 mM imidazole elution peak were identified as formate dehydrogenase (FDH). No CYPtr was identified.

Efforts were made to improve the yield of secreted CYPtr. To this end, colony 2 was further investigated in 2 ml culture experiments using 24-well deep-well plates. Using these milli-scale cultures allowed for testing multiple different additives and constructs to be performed in triplicate. In a first experiment, it was tested whether medium supplementation with either iron, δ-aminolevulinic acid (5-ALA), hemin or a combination of all three would increase CYPtr yield in the medium. Even though *P. pastoris* possesses the heme biosynthetic pathway and CYP intracellular production in *P. pastoris* has been accomplished without medium supplementation [328]–[333], the endogenous heme might not suffice to support the recombinant production and secretion of CYPtr. For example, for the recombinant production and secretion of another heme protein, heme peroxidase, in *P. pastoris*, the addition of hemin increased the yield of active enzyme in the medium [334]. Seven different conditions were tested and compared to the non-supplemented condition. Each of the three additives were added separately in two different concentrations, and a combination of the three additives was used in the seventh condition. The medium was collected after growth and recombinant CYPtr production and concentrated for subsequent western blot analysis (Figure 3.39). In each case, CYPtr was successfully secreted and indeed, more intense bands were observed although not to a large extent. Especially the supplementation with 0.5 mM 5-ALA and with a combination of the three additives consistently showed more intense bands across all three blots.

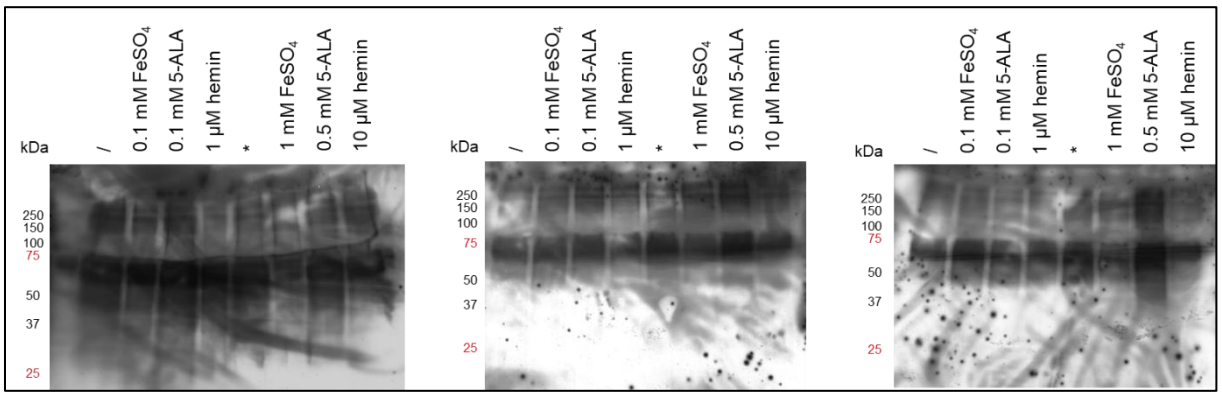

Figure 3.39: Western blot of DOC-TCA concentrated medium samples after milli-scale expression test in 24-well deep-well plates. Seven different medium supplementations were tested and compared to a non-supplemented culture. The three blots show the three biological replicates. * = 0.1 mM $FeSO_4$, 0.1 mM 5-ALA and 3 µM hemin.

It has been mentioned above that co-expression of CPR can markedly influence CYP production and stability. In parallel to the construction of a *cypntHis-cprntHis* co-expression vector, a *cyptr-cprntHis* co-expression vector was created. Eight colonies were picked upon transformation, of which only four contained both *cyptr* and *cprntHis*, based on colony PCR. These four colonies were subsequently grown for a small-scale expression test after which a medium sample was taken, concentrated and analysed by western blot. It was seen that CYPtr was present in the medium in a higher yield after CPRntHis co-expression (Figure 3.40A). CPRntHis indeed was produced, as shown by western blot of the isolated microsomes (Figure 3.40B). Of note, CYPtr contains an N-terminal Histag, whereas the Histag of CPRntHis is present at the C-terminal end. The anti-His antibody used for CPRntHis detection by western blot is specific for the Histag at the C-terminus. The signal seen in the microsomes (Figure 3.40B) is thus expected to be specific for CPRntHis and should not give a false positive due to CYPtr retained in the microsomes. The MW of CPRnt is much higher than of CYPtr (both enzymes differ in MW by 20 kDa). However, they appear to be equal in MW on blot raising questions about the nature of the signal in the microsomal extract. It must be noted that the western blot marker is not suited for exact MW determination, especially in the higher MW region, as in this region the marker is often overdeveloped and consequently, individual marker bands become more difficult to distinguish.

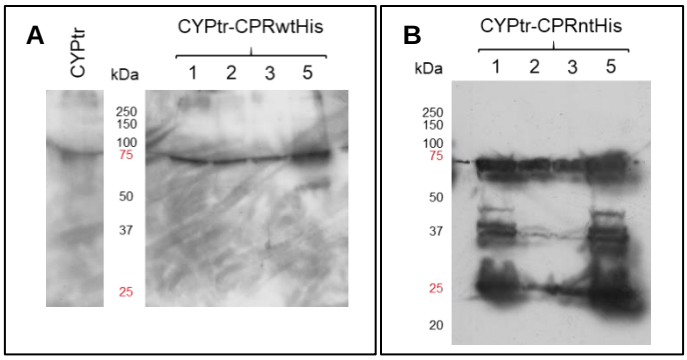

Figure 3.40: Western blot after small-scale expression test of four colonies picked after *cyptr-cprntHis* co-expressing vector transformation in *HAC1* co-expressing *P. pastoris*. A: Concentrated medium samples. The anti-His-HRP antibody (MA1-21315-HRP, Invitrogen) was used for CYPtr (58 kDa) detection. B: Microsomes. The anti-His(C-term)-HRP antibody (R931-25, Invitrogen), specific for C-terminal Histags, was used for CPRntHis (78 kDa) detection.

Further efforts for CYPtr yield improvement and purification optimizations were set aside. It was decided to focus on obtaining an active enzyme first because only then it becomes of interest to also optimize the production and purification process.

4 Discussion

4.1 The truncated CYP52A13 is not active

The original aim of this work was to create an enzyme reactor for the synthesis of long-chain α,ω-DCAs. Therefore, we designed constructs to deliver soluble forms of the CPR-a and CYP52A13 that could be introduced in such a reactor. Unfortunately, while we could achieve recombinant production of the proposed truncated form of CYP52A13, this did not result in an oleic acid-converting construct. This can be due to either an incompatibility with the redox partners used, or due to inherent inactivity of CYPtr.

Three alternative redox partners were considered for the delivery of electrons from NADPH to CYPtr; (1) the endogenous CPR from *P. pastoris*, (2) CPRtr, either recombinantly produced and added to the microsomes as a separate enzyme, or fused to CYPtr, and (3) BMR, fused to CYPtr.

In the first assays, the use of endogenous CPR from *P. pastoris* as a redox partner did not result in DCA formation, nor was NADPH consumed to a higher extent upon substrate addition. In fact, several examples were found in literature where the low level of endogenous CPR from *P. pastoris* was able to reduce the recombinant CYP [328]–[333]. However, other CYPs did not show activity when this CPR served as the only redox partner [335] and it is a generally used strategy to overexpress either the homologous or a heterologous CPR when CYP is recombinantly produced in *P. pastoris* [237], [331], [336]–[338]. Also fusion to the reductase domain of the self-sufficient CYP102A1 did not yield an active protein product. It is documented that changing the redox partner can have a marked influence on the catalytic properties of CYP [264]–[267]. Therefore, it was decided to focus further on using the native redox partner CPR-a.

In a first series of experiments, an attempt was done to produce CPRtr in a soluble and secreted form using the expression host *P. pastoris*. Indeed, CPRtr could be purified from the medium of the NRRL-Y-11430 strain by IMAC. Herein, secreted CPRtr proved to be active towards the nonphysiological redox partner cytochrome c. However, no increased NADPH consumption was seen upon substrate addition, indicating that the delivery of electrons to CYPtr was not accomplished. On the one hand, this can be ascribed to the low activities obtained. On the other hand, it must be considered that CPRtr might not be able to deliver electrons to CYPtr. It has been reported in literature that the binding site of CPR to either cytochrome c or CYP overlaps [202], negating the hypothesis that truncating CPR prevents interaction with its redox partner. However, we remind that the very first CPR crystal structure to be resolved, i.e. rat CPR obtained after a limited trypsin digest, was able to reduce cytochrome c but not CYP [191]. Therefore, it cannot be excluded that our CPRtr is unable to reduce our CYPtr. In the course of this PhD study, the structure of a N-terminally truncated CPR-b from *C. tropicalis* strain 1230, has been resolved [195]. This reductase shares 98.7 % amino acid sequence identity with CPR-a and the activity was similar. He and Chen inferred that the allelic *cpr* variants from *C. tropicalis* 1230 were the same as those from *C. viswanathii* ATCC 20336 [310]. Ebrecht et al showed that

this truncated CPR was able to not only reduce cytochrome c, but also CYP (more specifically CYP52A21 from *C. albicans*) [195]. It must be noted that they only deleted amino acids 2 to 22. One might thus wonder if the truncation applied in this dissertation (36 residues were deleted), led to an inability to interact with CYPtr. We based our design on the report of a truncated CPR from *S. cerevisiae*. In earlier reports of self-sufficient CYP construction, Shibata et al published a study where this CPR from *S. cerevisiae* was truncated differentially for its use in a CYP17A1 chimeric construct. They reported that the deletion of up to 53 amino acids all resulted in 90-95 % progesterone conversion, be it with different 17α-hydroxylase activities [269]. This strengthens the hypothesis that indeed our CPRtr cannot only reduce cytochrome c but CYPtr as well.

Throughout the expression experiments, it became evident that the CPRtr was very unstable, which is likely to be the result of removing the enzyme from its membrane environment. High yields could not be obtained and a high amount of intracellular degradation was observed as well, indicating folding issues and consequently directing the misfolded protein to the cytoplasm for degradation. Furthermore, using the recombinant host co-expressing *HAC1*, precipitation issues during enzyme purification occurred. In fact, the latter strain led to a significant yield increase and improved removal of the secretion signal. This was concluded based on two results. Firstly, it was observed that only one band appeared with the right MW on blot, whereas two bands were visible after CPRtr production in NRRL-Y-11430. The two bands differed only slightly in MW and it was thought that this resulted from an unprocessed signal sequence. Only one band appears upon *HAC1* co-expression, indicating that the signal sequence was processed correctly in that case, which might lead to increased secretion. Indeed, the co-expression of *HAC1* has been shown to aid in the signal sequence processing in efforts to produce the adenosine A2A receptor [315] and increased yield of secreted protein has been reported with this strain [339]. Secondly, a significant increase of the absorption peak eluting with 100 mM imidazole was observed in case where no precipitation occurred either during sample preparation preceding IMAC or during sample loading on the HisTrap column. CPRtr produced in the *HAC1* co-expressing strain could not be recovered in any case, and thereby proved to be highly unstable. These instability issues appeared to be more pronounced than in case of CPRtr production in NRRL-Y-11430. From this, it was hypothesized that inclusion of the cofactors FAD and FMN became more limiting, whereas in the wild type strain, the lower production rate can be better followed by cofactor inclusion. This hypothesis could theoretically be confirmed by measuring the FAD and FMN concentrations spectrophotometrically and compare these to the CPRtr concentration. The concentration of CPRtr in crude samples can be determined based on the cytochrome c reductase activity [340]. Alternatively, total protein concentration can be measured from a pure CPRtr sample. However, the enzyme was already lost during the first steps of enzyme collection and aforementioned assays are therefore not possible without further optimizations.

CPRtr was fused N-terminally to the C-terminus of CYPtr in an effort to create a self-sufficient CYP. By creating a chimer, it was no longer necessary to separately produce CPRtr, which proved to be difficult and where yields, too low for functional assays, were obtained. Clear cytochrome c reducing activity was observed by the CPRtr, present as the reductase domain of the created self-sufficient CYPtr. If the truncated reductase indeed is able to reduce CYPtr as well, as discussed above, than why was oleic acid not converted? Several possible answers can be given to this question. Uncoupling reactions always occur to some extent and these increase when for example the substrate does not fit the binding pocket well or alterations are made concerning the redox partner [176]. Theoretically, the construction of a self-sufficient enzyme should increase coupling efficiency by improving the electron transfer. However, coupling might not be sufficient in this chimer for efficient reduction of CYPtr. The possibility of uncoupling was addressed in this dissertation by assaying whether NADPH was increasingly consumed in presence of oleic acid compared to NADPH consumption in absence of oleic acid by microsomes after recombinant expression. If substrate binding occurs, followed by the delivery of electrons from NADPH to CYPtr via CPRtr, it was expected that an increased NADPH consumption would be observed. If increased NADPH consumption does not match the level of FA hydroxylation, uncoupling occurs through one of the three shunts (autooxidation, peroxide and/or oxidase shunt). However, no such increased NADPH consumption was seen. Taken together with the absence of hydroxylated products in the GC chromatograms, this led to the conclusion that no hydroxylation, nor uncoupling is happening. Next to these shunt pathways, other uncoupling reactions, independent of substrate binding, have been mentioned in literature [176]. For example, the reduced flavin cofactors of CPR could interact with molecular oxygen, whereby ROS are formed [195], possibly harmful for the recombinant CYP. In fact, such an assay was performed, measuring the formed hydrogen peroxide after complete oxidation of NADPH, and 3.54 % of the added NADPH resulted in hydrogen peroxide formation (data not shown). No excessive uncoupling reactions thus seem to happen. To finally address whether CPRtr can transfer electrons to CYP, one could exploit the fact that the reduced CO-bound form shows a characteristic absorption peak at a wavelength of 450 nm. Hayashi et al showed how the redox partner was able to reduce the heme group by observing an increased 450 nm absorption peak upon NADPH reduction in presence of CO and substrate [341]. Unfortunately, we did not have access to facilities to produce CO-bound CYPtr.

An aspect that has not been covered, is the involvement of cytochrome *b5*. Three possible contributions of this enzyme have been reported, i.e. cytochrome *b5* has an allosteric effect, it delivers the second electron to CYP or both electrons are delivered by cytochrome *b5* [139]. Functional assays performed by Eschenfeldt et al proved that CYP52A13 and its native redox partner CPR-a actively convert oleic acid when recombinantly produced in insect cells. No cytochrome *b5* was added or co-expressed [24]. Therefore, it was hypothesized that this third enzyme partner was not required for activity. However, addition of this enzyme or co-expression might still be valuable. For example it has been reported to positively influence the coupling efficiency of chimer constructs in several cases [271].

The fact that the reductase domains showed to be active, together with the absence of substrate-induced uncoupling, would lead to the conclusion that CYPtr is not active. It might be asked whether the truncated form is correctly folded and present as a holo-enzyme and if the substrate is able to bind in the active site in presence of the unprocessed secretion signal. These questions as well could be further addressed by means of spectrophotometric assays. Upon reduction of and CO binding to the heme-iron, the Soret peak of the substrate-free enzyme shifts from 418 nm to about 450 nm. This only occurs if cysteine thiolate is retained as the proximal ligand, thereby proving correct folding. Furthermore, substrate binding could be assessed by spectrophotometrically measuring the Soret peak shift occurring upon substrate binding. The Soret peak shifts to 390 nm upon conversion of the low spin state to the high spin state of the heme-iron. This happens when the substrate binds and thereby displaces the axial water ligand [177].

The location of CYPtr needs further consideration as the truncated enzyme contains an α-factor secretion signal, thereby directing the enzyme post-translationally to the ER for subsequent processing and secretion through COPII vesicles. However, no secretion was observed using *P. pastoris* NRRL-Y-11430 and the enzyme was found in the microsomal fraction. Normally, class II CYPs are anchored in the ER membrane via the N-terminal transmembrane helix at the cytosolic side. However, this anchor was removed in CYPtr and replaced by the α-factor secretion signal aiming for translocation inside the ER lumen. Is CYPtr retained at the cytosolic side? The protein is translocated post-translationally into the ER lumen. If the protein is already folded in the cytoplasm, this translocation is blocked and the CYPtr would thereby be present at the cytoplasmic side. A hydrophobic patch, consisting of the N-terminal part of the catalytic domain and the F-G loop, was reported to enable additional membrane interaction. As the F-G loop was kept intact due to involvement in substrate binding, membrane interaction might still occur through this hydrophobic patch. Alternatively, CYPtr might be retained in the ER lumen [342]. This would be an alternative reason why both the endogenous CPR from *P. pastoris* and the recombinant CPRtr would not be able to reduce CYPtr. Furthermore, glycosylation already starts in the ER lumen, which might be a reason for the seemingly higher MW observed in Figure 3.8. In case of the chimeric construct, it was hypothesized that the enzyme was not translocated across the ER membrane into the ER lumen. On the one hand, the chimer was retrieved in the soluble supernatant around the expected MW upon washing microsomes with a high salt buffer. On the other hand, reducing activity was observed for the chimeric construct upon addition of cytochrome c and NADPH. Based on the chimeric construct results, it was hypothesized that the lack of oleic acid conversion could be attributed to an inactive CYPtr enzyme.

The different questions that were asked throughout the discussion are shown in light of the catalytic mechanism in Figure 3.41.

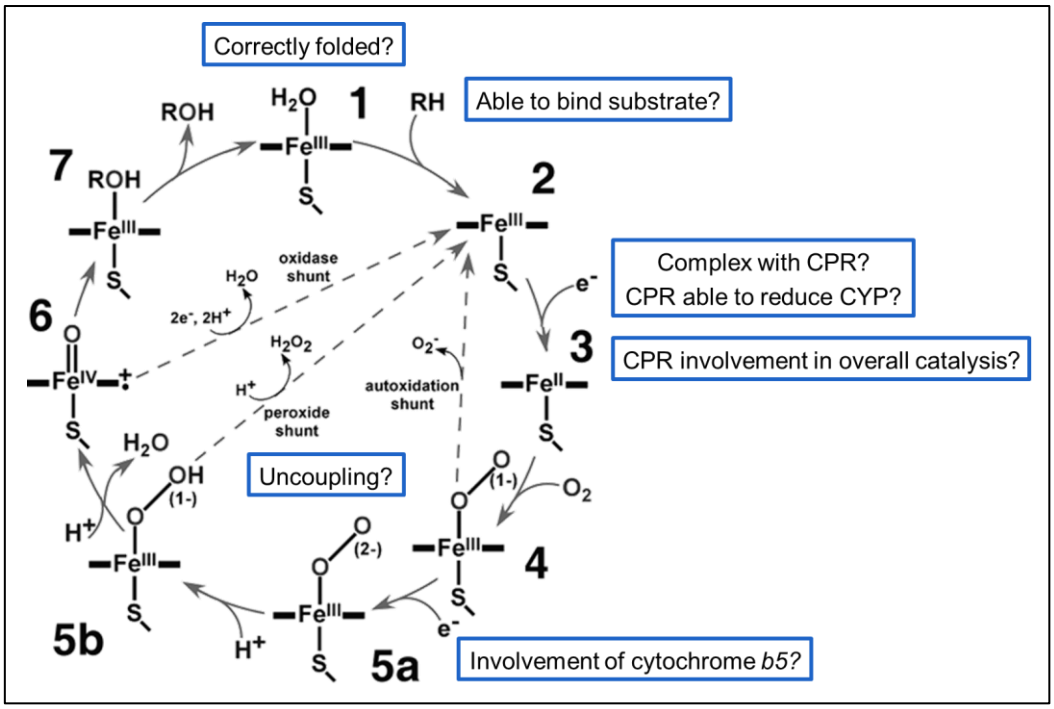

Figure 3.41: Catalytic mechanism including the different questions that were asked throughout this discussion.

4.2 The yield of CYP52A13 in a nontruncated form is too low for functional assays

The first assays showed that CYPtr did not actively convert oleic acid. To eliminate the question whether the truncations were the cause of this inactivity, it was decided to start producing nontruncated CYP52A13 in the native ER environment before testing alternative truncations. Using CYPnt, the intention was to test different linkers, and play with the amount of deleted residues in the CPR N-terminus, in order to acquire the self-sufficient CYP with the best possible coupling efficiency. However, it quickly became clear that CYPnt production in the heterologous host *P. pastoris* was not straightforward and only low yields were obtained. Eschenfeldt et al used insect cells for the recombinant production of CYP52A13 and CPR-a and they as well could not identify their recombinant proteins by SDS-PAGE analysis, visualized by Coomassie-staining [24]. Performing activity assays, they showed that only about 10 % oleic acid was converted to both OHFA and DCA. Therefore we pursued an increased CYPnt yield. Unfortunately, this was unsuccessful. Even though yeasts, being eukaryotes, possess the ER required for class II CYPs, it thus still proves to be advantageous to modify the N-terminus. Indeed, Schoch et al and Galanie et al reported that modifying the N-terminus of class II CYPs improved recombinant production in yeast [237], [238]. Schoch et al reported that this could be ascribed to the limiting space in the ER membrane. Galanie et al observed that their CYP enzyme (SalSyn) was present as three forms with different MWs. This was ascribed to the incorrect processing of the N-terminus, thereby orientating the CYP towards the ER lumen, amendable for glycosylation. Altering the N-terminus, enabled correct processing and a higher activity was accomplished. In this dissertation, an alternative N-terminal artificial sequence is thereby provided to increase the CYP yield in *P. pastoris*, i.e. the N-terminal tag, described by Lamb et al for the crystallization of CPR from *S. cerevisiae* [193]. Due to the fact that higher yields were obtained with this tag but the construct did not prove to be active,

alternative truncations might be tested. It must be noted that the CYP yield can differ from one strain to another, as well as differ between different recombinant hosts. For example, Hausjell et al demonstrated that the highest yield for the plant CYP chalcone 3-hydroxylase was obtained with the *P. pastoris* MutS strain KM71H, compared to Mut+ strains GS115 and SMD1168H. It was hypothesized that this difference could either be ascribed to a reduced stress, due to the fact that no high levels of AOX1 are being produced, or due to the reduced formation of hydrogen peroxide. Also the two Mut+ strains showed different expression yields [234] and CYPs have been successfully produced in Mut+ strains as well [328], so both MutS and alternative Mut+ strains might be tested in parallel in order to look for the best producing strain.

In order to obtain a positive control, as well as to increase CYPnt yield, a co-expression construct was created. Unfortunately, the results indicated a large imbalance between CYPntHis and CPRntHis production, with CPRntHis being produced to a much higher level and showed to decrease instead of increase CYPntHis yields. It has been stated that CPR as well can attribute to uncoupling by delivering the electrons from NADPH to molecular oxygen instead of CYP. The high level of CPRntHis might thus lead to increased ROS production, possibly inactivating CYPntHis. Indeed it has been reported that it is advantageous to express *cpr* from a weak promotor [132]. As such, a positive control could not be created. Alternatively, microsomes from *C. viswanathii* ATCC 20962 were isolated after inducing the culture with oleic acid. However, it proved to be difficult to remove the oleic acid and formed DCA from the isolated microsomes, thereby interfering with subsequent activity assays. Indeed, alternative substrates could be used in order to induce expression of the CYP52A13 encoding gene, such as the alkane octadecane. However, it must be considered that these microsomes also contain other CYPs which are induced upon oleic acid addition, e.g. CYP52A17 [23]. For these reasons, attention first went to the recombinant CYPnt production in *P. pastoris*.

4.3 A first step towards the production of a soluble secreted class II CYP has been made

A remarkable result was the secretion of CYPtr in the medium when produced in the *HAC1* co-expressing *P. pastoris* strain. To our knowledge, this is the first report of a soluble and secreted class II CYP. Unfortunately, this was only obtained in small-scale experiments. After a midscale expression test and IMAC purification, no CYPtr was recovered. Instead, DAS2 and FDH were found in the first two elution peaks. Both enzymes are important catalysts of the methanol utilization pathway and can take up 20 % of the total protein content upon methanol induction [312]. DAS2 resides in the peroxisome whereas FDH is localized in the cytosol. Although the optical density of this midscale culture was not significantly lower than observed during other protein production experiments, this might still indicate a significant amount of cell lysis occurring during this midscale expression test. Further efforts are thus required in order to obtain soluble and secreted CYPtr in a sufficient yield for further applications. Clonal variations were seen and it appeared that colony 2 produced a more stable CYPtr than colony 1 and 3. This might be ascribed to differences in gene copy numbers or different gene integration events (although efficient homologous insertion in the *AOX1* locus is targeted due to digestion in the AOX1 promotor). However, clonal variations are not solely attributed to these factors

[343]. Therefore, it might be interesting to investigate differentially expressed genes between the three initial colonies.

Whether CYPtr produced as a soluble and secreted enzyme can be used as a biocatalyst, is questionable. Whereas no degradation was observed in case of CYPtr production in NRRL-Y-11430, where the enzyme was retained in its natural membrane environment of the ER, significant degradation was seen upon *HAC1* co-expression and the CYPtr could not be purified. Removing the enzyme from its membrane environment proved to be a hard to overcome issue for further biocatalyst development. However, the N-terminal tag improved yields and increased solubility as it enabled secretion. Although the chimer CYPtrCPRtr was not secreted using the same strain, it could be retrieved in a soluble form using a high salt buffer and this in absence of detergents, further highlighting the value of this artificial N-terminal tag.

5 Conclusion

CYPs are very interesting enzymes for their large potential in industrial biocatalysis. Unfortunately, they are notorious for their instability and low activities. The class II CYPs come with an extra challenge, i.e. they are hydrophobic membrane-bound enzymes. CYP52A13, of interest in the production of long-chain α,ω-DCAs, as well proves to be difficult to produce recombinantly. Only low yields were obtained in the yeast *P. pastoris* and activity towards the substrate oleic acid could not be confirmed. Even so, an N-terminal modification was proposed, leading to increased yields in the strain NRRL-Y-11430, solubilization in absence of detergent and secretion to the medium in the *HAC1* co-expressing strain.

Chapter 4
Outer membrane vesicle display of a self-sufficient CYP as a nanobiocatalyst

Delphine Devriese[1], Pieter Surmont[2], Frederic Lynen[2], Bart Devreese[1]

[1]Laboratory of Microbiology – Protein Research Unit, Ghent University
[2]Separation Science Group, Ghent University

Authors contribution: Delphine Devriese wrote the whole chapter and performed all experiments. Pieter Surmont performed GC-MS of the silylated samples. Frederic Lynen is the supervisor of Pieter Surmont. Bart Devreese is the supervisor of the author.

The part of this chapter describing CYP102A1 display on *E. coli*-derived OMVs as a proof-of-concept is in preparation for publication.

Abstract

The cytochrome P450 (CYP) monooxygenase superfamily has large potential in synthetic biology applications. For example, several CYPs have been shown to regioselectively oxidize fatty acids at their ω-position, such as the *Candida viswanathii* CYP52A13. This is of interest in the production of long-chain α,ω-dicarboxylic acids (DCA), as only a limited amount of long-chain DCAs is commercially available and the price is elevated. CYPs are generally applied in fermentation processes, due to their low stability and the need for a redox partner and cofactor. Unfortunately, substrate uptake and/or product transport limitations are frequently encountered and side reactions occur due to other enzymes in the cellular environment. The production of a solubilized CYP52A13 was pursued for its use in an *in vitro* biocatalytic process for long-chain DCA production. However, this proved to be very challenging and no success was booked so far. Therefore, another route was taken, i.e. the display of a self-sufficient CYP on the surface of outer membrane vesicles (OMVs), derived from *Escherichia coli*. This strategy provides an *in vivo* immobilization technique, possibly enhancing stability of CYP enzymes. Using the self-sufficient enzyme CYP102A1 from *B. megaterium*, a proof-of-concept could be delivered. Unfortunately, no DCA-producing catalyst was obtained.

1 Introduction

Previous chapter showed that recombinant production of class II-type CYPs is not straightforward and requires a lot of time and effort for optimization. Producing CYP in a soluble form for industrial application is very challenging and no great successes have been booked so far. Therefore, an alternative approach for the application of class II-type CYPs in industrial catalysis was investigated, i.e. bacterial surface display and display on outer membrane vesicles (OMVs). Surface display offers some advantages over both fermentation and *in vitro* biocatalysis. By displaying the enzyme on the surface, substrate uptake and product transport limitations, which is a major hurdle in the production of long-chain α,ω-DCAs by fermentation, are circumvented. Moreover, the displayed enzyme is immobilized *in vivo* on the outer membrane, creating a membrane environment which could enhance stability of CYPs from class II.

CYP102A1, the natural self-sufficient class VIII enzyme, has been displayed on the surface of *E. coli* using three different approaches. In a first publication, Yim et al used the ice-nucleation protein (INP) InaK from *Pseudomonas syringae,* in a truncated form [250]. INPs are found in plant pathogenic bacteria and are composed of three distinct domains: a N-terminal unique domain containing the membrane anchor, a C-terminal unique domain and a central repeating domain [251]. This central domain serves as a template for ice nucleation and was deleted, as this is not required for surface display [250]. In a second publication, Ströhle et al reported the display of CYP102A1 by Autodisplay. The so-called autotransporter AIDA-I was used. This protein consists of an adhesin passenger with a translocator domain at the C-terminus, fused by a linker peptide. Ströhle et al displayed CYP102A1 by fusing the AIDA-I translocator (including the linker peptide) to its C-terminus and clone the corresponding nucleotide sequence in frame of the gIII signal peptide-encoding sequence from the pBAD gIII vector [254]. Thirdly, a very recent article reports the display of CYP102A1 by *in situ* SpyCatcher-SpyTag interaction. The Spycatcher protein was fused to the Lpp-OmpA hybrid surface display anchor. CYP102A1 was fused to the SpyTag peptide. Both proteins were expressed separately in *E. coli*, after which the SpyCatcher protein and SpyTag peptide form a covalent isopeptide bond, thereby coupling CYP102A1 to a membrane anchor posttranslationally, and enabling translocation to the surface of *E. coli* [255]. The surface display of class II-type CYPs on the surface of *E. coli* has been reported as well, i.e. human CYP3A4 and human CYP1A2. This was done by the group of Jose by means of Autodisplay [247], [249].

Gram-negative bacteria are known to produce OMVs. These are spherical particles of 20-250 nm, derived from the outer membrane (OM). Numerous functions have been ascribed to these OMVs such as biofilm formation, pathogenesis and communication [344]. In our lab, OMVs have been investigated in frame of the multidrug-resistant *Stenotrophomonas maltophilia* [345]–[347]. OMVs have emerged as a valuable vaccine delivery vehicle as they are able to modulate the immune response and much research has already been done to engineer these OMVs in order to obtain fit-for-purpose vesicles [348]. Next to vaccine delivery, OMVs have been suggested to serve as synthetic nanoreactors. Park et al were able to display a scaffold

protein of 107 kDa on the surface of OMVs derived from *E. coli*, using an INP of *P. syringae*. Several enzymes were assembled on this scaffold, resulting in OMVs able to hydrolyse cellulose. The display on OMVs led to a marked activity increase when compared to the same scaffold, surface displayed on yeast cells [349]. It was hypothesized that this was due to the nanoscale dimension of the OMVs, leading to a higher enzyme:volume ratio and improved substrate accessibility, in contrast to the microscale dimension of a yeast cell [350]. Therefore, this approach caught our attention.

In normal conditions, *E. coli* produces only a low amount of OMVs. For the use of OMVs as a biocatalyst, inducing a hypervesiculating phenotype is preferred [350]. OMVs are formed by blebbing of the OM. This blebbing is thought to occur at regions where crosslinks are reduced, leading to decreased envelope stability (Figure 4.1) [344], [351], [352]. The Tol-Pal complex is composed of five proteins, that is TolA, TolB, TolQ, TolR and Pal, and spans the entire *E. coli* envelope. Non-covalent interactions between the peptidoglycan (PG) layer and the Tol-Pal complex serves as one of the crosslinks, providing envelope stability. Deleting one or more of these Tol-Pal complex-encoding genes reduces the amount of crosslinks between the PG and the OM, consequently leading to increased vesiculation [353]. Next to the Tol-Pal complex, the deletion of other genes, encoding proteins involved in envelope crosslink modulation, were reported to result in a hypervesiculating strain, for example, knocking out the gene encoding the OM lipoprotein NlpI. It was hypothesized that this deletion resulted in a PG turnover imbalance. Proper crosslinking was therefore prevented, consequently leading to hypervesiculation [354], [355]. Not only reduced crosslinking, but also envelope stress was reported to be involved in OMV biogenesis (Figure 4.1). Disruption of genes involved in the σ^E stress response pathway, that is *degS, degP and rseA*, led to a marked OMV increase. This pathway is activated in response to misfolded protein in the periplasm and it was hypothesized that OMV formation was a result of protein accumulation, leading to bulging of the OM and finally secreting these accumulated proteins through OMVs [356]. Furthermore, OMV biogenesis occurs as a response to external factors such as temperature, nutrient availability and quorum sensing, where the quorum sensing molecule *Pseudomonas* quinolone signal (PQS) intercalates in the OM, thereby inducing membrane curvature and ultimately leading to OMV production (Figure 4.1) [344], [351], [352]. Several antibiotics as well have been shown to induce OMV production. Subinhibitory concentrations of antimicrobial peptides which intercalate with the membrane, such as polymyxin B and colistin, lead to increased vesiculation, where the OMVs serve as a protective response [357]. Similar phenomena were observed when using β-lactam antibiotics, which blocks PG synthesis and thereby impairing the cell wall integrity. Upon addition of β-lactam antibiotics to resistant bacteria, β-lactamase-containing OMVs are produced [346], [358]. In this context it is worth noting that also molecules inducing the SOS response (e.g. the antibiotic ciprofloxacin), have been reported to induce vesiculation, although not OMVs but outer inner membrane vesicles (OIMV) are formed as a result of explosive cell lysis [347], [352].

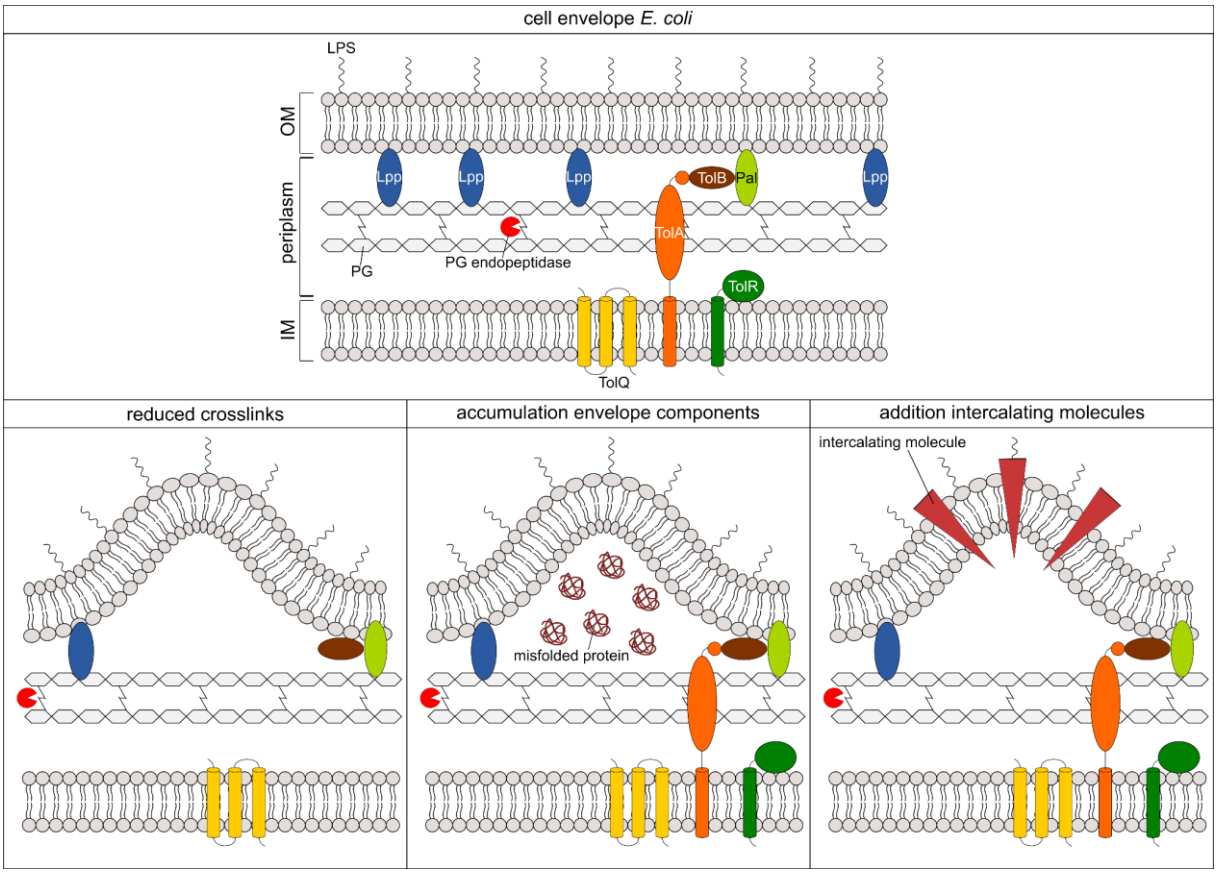

Figure 4.1: (top) Cell envelope of the gram-negative bacterium *E. coli*. The envelope contains two membranes, i.e. the outer and inner membrane (OM and IM, respectively), with the periplasm, containing a peptidoglycan layer (PG), in between. The OM consists of an asymmetrical bilayered membrane, existing of a phospholipid inner leaflet and an outer leaflet containing lipopolysaccharides (LPS). Envelope stability is enabled through several crosslinks such as the covalent crosslink of Braun's lipoprotein (Lpp) and the non-covalent interactions between the PG layer and the Tol-Pal complex (TolA, TolB, TolQ, TolR and Pal). (bottom) Several routes leading to outer-membrane vesicle biogenesis. Deletion of TolA- and TolR-encoding genes reduces the amount of crosslinks, leading to hypervesiculation. Accumulation of misfolded protein leads to bulging of the OM and finally to secretion of these accumulated proteins through OMVs. Molecules, intercalating in the OM, induce membrane curvature and lead to OMV production.

In this chapter, the surface display of a self-sufficient CYP, able to hydroxylate FAs at the ω-position, was investigated. Using CYP102A1, it was explored if the surface display could be taken one step further, i.e. if OMVs displaying CYP102A1 could be isolated. The INPs and Autodisplay have been put forward as the systems of choice for the display of complex and large proteins [252], [253], exemplified by their application in CYP display. The use of an INP was chosen for surface and OMV display in this dissertation, in analogy with the surface display of CYP102A1 by Yim et al [250] and the OMV display of INP-Scaf3 (107 kDa) by Park et al [349]. It was opted to display the two redox partners as one self-sufficient fusion enzyme, in an attempt to mimic the highly efficient self-sufficient CYP102A1. Additionally, only one enzyme needs to be translocated to the OM instead of two. When expressing two or more proteins using the same translocation pathway, competition occurs, reducing expression [252], [349]. Several hypervesiculating mutants, as well as polymyxin B induced strains, were tested for their ability to produce OMVs, displaying a self-sufficient CYP.

2 Material and methods

2.1 Materials

All chemicals were purchased from Sigma-Aldrich, unless specified otherwise.

2.2 Strains and media

For cloning as well as for surface display, the *E. coli* JM109 strain (New England Biolabs (NEB)) was used. Cells were grown in Luria-Bertani (LB) medium. LB consists of 1 % (w/v) NaCl (Merck), 0.5 % (w/v) yeast extract (Lab M) and 1 % (w/v) tryptone (Lab M). For agar plates, 1.5 % (w/v) agar (Lab M) was included. 100 µg/ml carbenicillin (Cb) (Gold Biotechnology) was added to medium and agar plates for selection of transformants and for plasmid maintenance during protein production.

For OMV display, several deletion mutants were used next to *E. coli* JM109 and are shown in Table 4.1. Two of the deletion mutants were ordered from the Keio collection [359]. In this collection, a duplicate of each deletion mutant is present. They included two independent mutants due to the fact that handling errors, cross-contamination and accumulation of secondary mutations can occur in high-throughput experiments. Therefore, both replicates for each deletion mutant was ordered for evaluation. For cloning, precultures and OMV quantification, cells were grown in LB medium. Strains from the Keio collection were grown in the presence of 25 µg/ml kanamycin (Kan) (Labconsult). In case of protein production, LB medium or Terrific Broth (TB) medium was used. TB consists of 1.2 % (w/v) tryptone (Lab M), 2.4 % (w/v) yeast extract (Lab M), 0.4 % (v/v) glycerol and phosphate buffer (0.017 M KH_2PO_4, 0.072 M K_2HPO_4). Alternatively, precultures were grown in non-inducing medium MDAG-135 and protein production was performed in autoinduction medium ZYM-5052, as described by Studier et al [360]. Here as well, 100 µg/ml Cb (Gold Biotechnology) was added to medium and agar plates for selection of transformants and for plasmid maintenance during protein production.

Table 4.1: *E. coli* deletion mutants, used for OMV display.

Strain	Deletion	Source
tolA 837	Δ*tolA*	*E. coli* Keio Knockout Collection Horizon Discovery: OEC4987-200826250
tolA 838	Δ*tolA*	*E. coli* Keio Knockout Collection Horizon Discovery: OEC4987-213606638
nlpI 863	Δ*nlpI*	*E. coli* Keio Knockout Collection Horizon Discovery: OEC4987-200828301
nlpI 864	Δ*nlpI*	*E. coli* Keio Knockout Collection Horizon Discovery: OEC4987-213605299
JC8031	Δ*tolRA*	Described in [353] Kindly provided by Prof. Dr. Laetitia Houot (Laboratoire d'Ingénierie des Systèmes Macromoléculaires, CNRS)

2.3 DNA constructs and molecular cloning

InaK was truncated for its use as a surface display anchor, and is from here on now denoted as InaKtr. A construct was designed, consisting of the N- and C-terminal unique domains and a C-terminal Asp-Pro-Gly linker, as described by Yim et al [250]. Truncation of the central domain was performed based on the previous report of Shimazu et al [361], where they deleted the central domain except for subdomain 1 (an incomplete 32-residue repeat, which may serve as a transition between the N-terminal and the repeating domain) and one 48 residue repeat from subdomain 4. Furthermore, a NheI restriction site and sequence encoding a Histag of six His residues at the C-terminus were introduced (Figure 4.2). This construct was codon optimized, synthesized and cloned in the vector pMAL-c4x by GenScript. Cloning was performed in between restriction sites NdeI and EcoRI, thereby excluding the maltose binding protein encoding gene. The pMAL-c4x vector was chosen for its *tac* promotor, as this promotor was proposed to be optimal for recombinant CYP expression in *E. coli* [229], [230]. The construct will further be referred to as InaKtr_pMAL (see Addendum).

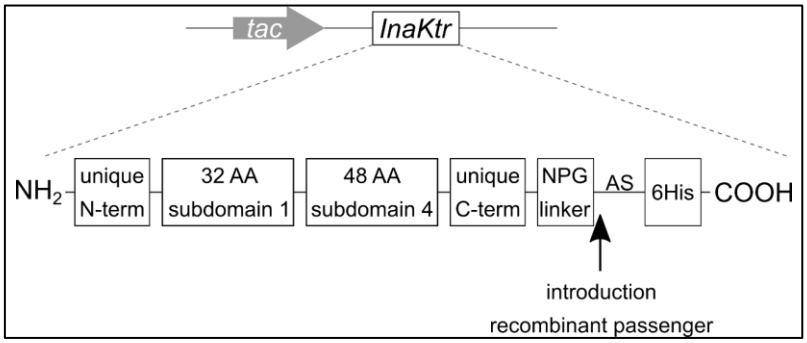

Figure 4.2: Design of the truncated InaK surface display anchor. The gene encoding InaKtr was under the control of the *tac promotor*. InaKtr consists of the unique N-terminal and C-terminal domain. Subdomain 1 and one 48 residue repeat from subdomain 4 were kept. At the C-terminus, a Asp-Pro-Gly linker was introduced, followed by a NheI restriction site, encoding Ala and Ser, and a Histag. The recombinant passengers are introduced in between the Asp-Pro-Gly linker and the Histag.

For all molecular cloning experiments, Circular Polymerase Extension Cloning (CPEC), was applied as described in previous chapter (Chapter 3 - 2.4) and all primers are given in Table 4.2. The gene encoding CYP102A1, otherwise denoted as BM3, was amplified from genomic DNA of *Bacillus megaterium* (kindly provided by Prof. Inge Van Bogaert (CSB, Ghent University)), using the primer pair ForBM3_CPEC InaKtr - RevBM3_CPECInaKtr. The InaKtr_pMAL vector was linearized and amplified using the primers ForInaKtr_CPECBM3 and RevInaKtr_CPECBM3. The amplified *bm3* gene was subsequently inserted in this linearized InaKtr_pMAL vector, in frame of the sequence encoding the Asp-Pro-Gly linker, preceding the NheI restriction site and the Histag encoding sequence, resulting in BM3 with an N-terminal InaKtr surface display anchor and a C-terminal Histag. The construct will be denoted as InaKtrBM3_pMAL. The final polypeptide chain is shown schematically in Figure 4.3.

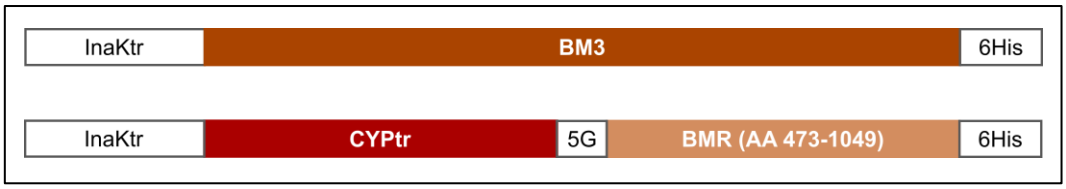

Figure 4.3: Schematic representation of polypeptide chains for the surface and OMV display of (top) the natural self-sufficient CYP102A1 from *Bacillus megaterium*, also referred to as BM3 and (bottom) the artificial chimeric construct of truncated CYP52A13, fused to the reductase domain of BM3, linked by a flexible linker of five glycine residues.

In order to place the expression of *inaKtrBM3* under the control of the *araBAD* promotor, the complete coding sequence was amplified with the primers ForInaKtrBM3_CPECpBAD and RevInaKtrBM3_CPECpBAD () for insertion in the vector pBAD/Myc-HisA. The vector was amplified and linearized using the primers ForpBAD_CPECInaKtrBM3 and RevpBAD_CPECInaKtrBM3 ().

For the surface display of a self-sufficient CYP52A13, the heme domain of BM3 was replaced by CYPtr (as described in Chapter 3 - 2.3) without the N-terminal tag. The native BM3 linker between the heme and reductase domain was replaced by a Gly-rich linker, in analogy with the optimized linker used for the so-called molecular lego approach as described by Degregorio et al [274] and Castrignano et al [275]. The insert gene encoding CYPtr was amplified from CYPtr_pPICZαB with the primers ForCYPtr_CPECInaKtr-BMR and RevCYPtr_CPECInaKtr-BMR (Table 4.2). The InaKtrBM3_pMAL construct was used as the template for vector amplification using the primer pair ForBMR_5G - RevInaKtrBM3_CPEC CYPtr (Table 4.2). Subsequent CPEC reaction resulted in the construct encoding the enzyme denoted as InaKtrCYPtrBMR. The final polypeptide chain is shown schematically in Figure 4.3. In a similar fashion, chimers were constructed where CYP52A13 was truncated differentially, and the design of these alternative truncations is described below. The respective *cyp* insert sequences were amplified from CYPnt_pPICZA, using the forward primers ForCYPΔ32_CPECInaKtr-BMR, ForCYPΔ25_CPECInaKtr-BMR and ForCYPnt_ CPECInaKtr-BMR, with the reverse primer RevCYP_AltTrunc (Table 4.2). The respective vector sequences were amplified using the forward primer ForBMR_5G and the reverse primers RevInaKtrBM3_CPECCYPΔ32, RevInaKtrBM3_CPECCYPΔ25 and RevInaKtrBM3_ CPECCYPnt (Table 4.2). CPEC then resulted in the constructs encoding the enzymes denoted as InaKtrCYPΔ32BMR, InaKtrCYPΔ25BMR and InaKtrCYPntBMR, respectively.

Table 4.2: Primers used for molecular cloning of the different expression constructs, by means of CPEC. The bold sequence indicates the overlapping region between vector and insert. The hyphen separates the non-annealing part of the primer from the sequence annealing to the template for initial amplification of the respective inserts and linearization of the respective vectors.

Primer	Sequence (5'-3')
ForInaKtr_CPEC BM3	**CAAAAGACGTGTGGGCTGGG-**GCTAGCCATCATCACCATCATCAC
RevInaKtr_CPEC BM3	**GCTGAGGCATTTCTTTAATTGTCAT-**GCCCGGGTCCTTAACTTCAATCC
ForBM3_CPEC InaKtr	**ATGACAATTAAAGAAATGCCTCAGC**
RevBM3_CPEC InaKtr	**CCCAGCCCACACGTCTTTTG**
ForInaKtrBM3_CPEC pBAD	**ATGACCCTGGATAAGGCG**
RevInaKBM3_CPEC pBAD	**GGCCCAAGCTT-CGAATTC**TTAGTGATGATGGTGATG
ForpBAD_CPEC InaKtrBM3	**GAATTCGAAGCTTGGGCC**
RevpBAD_CPEC InaKtrBM3	**CGCCTTATCCAGGGT-CAT**GGTTAATTCCTCCTGTTAGCC
ForBMR_5G	**CCTAGGTCAGGCGGTGGCGGTG-GC**AAAAAGGCAGAAAACGCTCATAAT
RevInaKtrBM3_CPEC CYPtr	**GTAACAGGCTTTGCACCAAGTT-T**GCCCGGGTCCTTAACTTCAATCC
ForCYPtr_CPEC InaKtr-BMR	**AAACTTGGTGCAAAGCCTGTTAC**
RevCYPtr_CPEC InaKtr- BMR	**GCCACCGCCACCGCCTGACCTAGG-**ATACATTTCAATATTGGCACCATCG
ForCYPΔ32_CPEC InaKtr-BMR	**GGTAGATACCTGATGTACAAGTTGGG**
ForCYPΔ25_CPEC InaKtr-BMR	**AGAGTCCTTGACTACTTCTACGG**
ForCYPΔnt_CPEC InaKtr-BMR	**ATGACCGTCCACGACATCATC**
RevInaKtrBM3_CPEC CYPΔ32	**CCCAACTTGTACATCAGGTATCTACC-**GCCCGGGTCCTTAACTTCAATCC
RevInatrKBM3_CPEC CYPΔ25	**CCGTAGAAGTAGTCAAGGACTC-T**GCCCGGGTCCTTAACTTCAATCC
RevInaKtrBM3_CPEC CYPnt	**GATGATGTCGTGGACG-GTCAT**GCCCGGGTCCTTAAC
RevCYP_AltTrunc	**GCCACCGCCTGACCTAGG-**GTACATTTCAATATTAGCACCGTC

To design the alternative truncations of CYP52A13, the Phyre2 [317] secondary structure prediction was investigated. According to this prediction, not the full N-terminal helix is incorporated in the membrane (Figure 4.4). Therefore, a first alternative was to only truncate the transmembrane helix up to AA 32 (indicated by the arrow in Figure 4.4) instead of the first 38 AAs, and no internal AAs were substituted. This construct will be referred to as CYPΔ32. Upon further investigation of the structure using Phyre2 in combination with additional literature study, another construct was considered. Monk et al described the first CYP structure without N-terminal truncations and with the transmembrane helix structure intact, i.e. *S. cerevisiae* lanosterol 14α-demethylase from the CYP51 family [165]. There it was seen that beyond Pro38, the transmembrane helix makes extensive polar contacts with the catalytic

domain. Also in other CYP structures, a proline-rich region was observed and truncation of this proline-rich domain resulted in loss of function [256]. Although CYP52A13 lacks such a proline-rich domain, this was investigated by looking at the alignment of CYP52A13 and CYP51 in Phyre2 (Figure 4.5). Here Pro38 aligned with residue 26 of CYP52A13 so it was decided to include a construct only lacking the first 25 residues, further denoted as CYPΔ25. Lastly, CYPs have been expressed in *E. coli* without an N-terminally truncated transmembrane helix. These constructs were incorporated in the membrane by providing a signal sequence from PelB or OmpA. After processing of the respective signal sequences, they showed that the native CYP was left behind in the membrane [242]. With this in mind, it was asked whether truncation of the N-terminal helix was necessary when providing the surface display anchor InaKtr. This construct will be referred to as CYPnt.

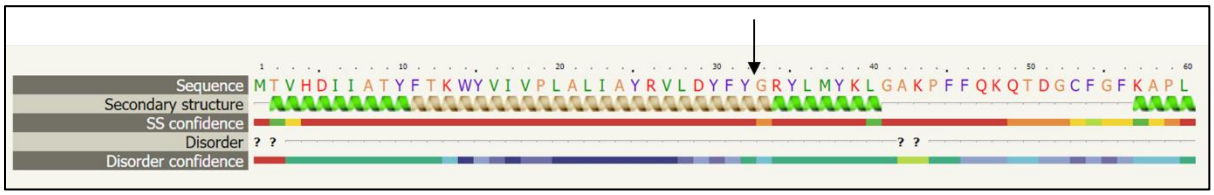

Figure 4.4: N-terminus CYP52A13 – secondary structure and disorder prediction using Phyre2. The arrow indicates the site of truncation in case of the Δ32 construct.

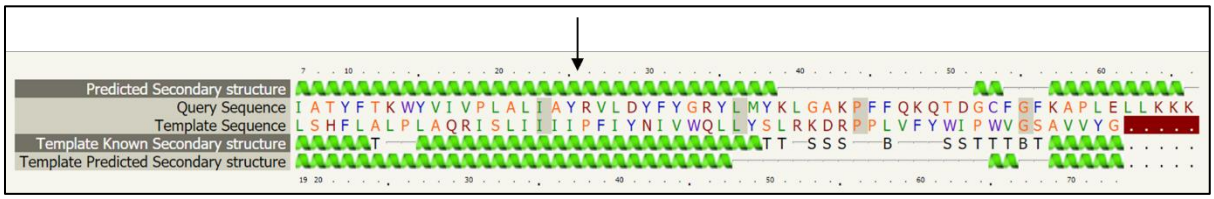

Figure 4.5: Alignment of the N-terminus of CYP52A13 (= query sequence) with lanosterol 14α-demethylase (= template sequence) using Phyre2. The arrow indicates the site of truncation in case of the Δ25 construct.

2.4 Culture conditions

All cultures were started with a 5 ml preculture using LB medium including the appropriate antibiotics in a 50 ml shake flask, which were incubated overnight at 37 °C and 200 rpm. In case of the OMV experiments, 1 % glucose was added for the pMAL constructs.

In case of surface display, *E. coli* JM109 was grown in 50 ml LB including 0.5 mM δ-aminolevulinic acid (5-ALA) and 100 µg/ml Cb, inoculated with 500 µl preculture, in a 500 ml shake flask at 28 °C and 200 rpm until an OD_{600} of 0.5. At this point, expression was induced by the addition of isopropyl-β-D-thiogalactoside (IPTG) (Gold Biotechnology) to a final concentration of 1 mM. The expression phase was carried out overnight and cells were collected the next day for subsequent assays and analyses.

For OMV quantification, the respective *E. coli* strains were grown in 50 ml LB (including Kan in case of strains from the Keio collection), in a 500 ml shake flask after inoculation with 500 µl preculture. The cultures were incubated at 37 °C and 200 rpm until an OD_{600} of 0.5, at which point the supernatant was collected for sOMV isolation as described below. Alternatively, cultures were grown overnight for sOMV collection in the stationary phase. In case of quantifying the polymyxin B induced hypervesiculation, JM109 was grown in 100 ml LB in a 1 l shake flask at 37 °C and 200 rpm until an OD_{600} of 0.5. The culture was split in two and polymyxin B (Labconsult) was added to a final concentration of 0.75 µg/ml. Incubation continued for another 3 h and sOMVs were collected.

When OMV display was pursued, the respective *E. coli* strains were grown in 100 ml LB or TB with 0.5 mM 5-ALA and 100 µg/ml Cb in a 1 l shake flask. For strains from the Keio collection, Kan was added and for the pMAL constructs, 0.5 % glucose was included. Cultures were incubated at 37 °C and 200 rpm until an OD_{600} of 0.5. At this point, the temperature was lowered to 28 °C and expression was induced either by IPTG addition to a final concentration of 0.3 mM (pMAL constructs) or by arabinose addition to a final concentration of 0.2 % (w/v) (pBAD constructs). Expression was carried out to the late exponential phase. In case of expression in *E. coli nlpI* 863, an overnight expression phase was performed. The cells and/or supernatant were subsequently collected for OMV collection. In early experiments, expression was induced with 1 mM IPTG, as indicated in the results section. No glucose was added in these cultures. Other expression times have been used, as indicated in the results section. In those experiments, the growth phase was carried out at 28 °C as well.

2.5 Cell lysis and outer membrane collection

Cells were collected by centrifugation at 20 000 g and 4 °C for 15 min. The supernatant was discarded and the cell pellet was resuspended in 50 mM Tris-HCl, pH 7.4, including protease inhibitor. Cell lysis was done by sonication on ice. Following lysis, the cell debris and unlysed cells were pelleted by centrifugation. A second sonication step was included and the two supernatant lysate samples were pooled for membrane collection by ultracentrifugation at 100 000 g and 4 °C for 1 h. The membrane pellet was washed with the same buffer, including 300 mM NaCl (Merck), and ultracentrifugation was repeated. Collection of the outer membrane (OM) was performed as described by Sandrini et al [362], with modifications. To solubilize the inner membrane (IM), the washed membrane pellet was incubated with 50 mM Tris-HCl, pH 7.4, including 2 % Triton X-100 for 30 min at room temperature with occasional mixing. A last ultracentrifugation step at 100 000 g and 4 °C for 1 h was performed, resulting in an OM pellet and solubilized IM in the supernatant.

2.6 OMV collection

2.6.1 sOMV collection

Spontaneously formed outer membrane vesicles (sOMVs) were isolated from the culture supernatant after pelleting the cells by centrifugation at 10 000 g and 4 °C for 10 min. The collected supernatant was filtered using a pore size of 0.22 µm. In later experiments, a switch was made to a 0.45 µm filter, as indicated in the results section. Next, the filtered medium was ultracentrifuged for 3 h at 100 000 g and 4 °C. The OMV pellet was washed with phosphate buffered saline (PBS), followed by ultracentrifugation. The final OMV pellet was resuspended in PBS for subsequent enzymatic assays and/or analyses. For OMV quantification, no washing step was included and the obtained pellet was immediately resuspended in 1 ml PBS. In case of OptiPrep™ density ultracentrifugation, the OMVs were also not washed and directly resuspended in a 45 % OptiPrep™ solution.

2.6.2 eOMV extraction

For EDTA extraction of outer membrane vesicles (eOMVs), the method described by Van de Waterbeemd et al [363] was used, with modifications. Cells were collected by centrifugation at 10 000 g and 4 °C for 10 min. The cell pellet was resuspended in 7.5 volumes (ml/g wet weight) 100 mM Tris-HCl, pH 8.6, 10 mM EDTA, and incubated for 30 min at room temperature. Cells were pelleted by centrifugation and the supernatant was collected. Subsequently, filtration was performed applying a pore size of 0.22 µm. Alternatively, a pore size of 0.45 µm was chosen, as indicated in the results section. eOMVs from the supernatant were pelleted by ultracentrifugation for 3 h at 100 000 g and 4 °C. Similar as for the sOMVs, wash and final resuspension was performed in PBS. In case of OptiPrep™ density ultracentrifugation, the OMVs were not washed and immediately resuspended in a 45 % OptiPrep™ solution.

2.7 SDS-PAGE and western blot

The same protocol as described in previous chapter (Chapter 3 - 2.9) was applied. Proteins were separated either using a 7.5 % or a 12 % polyacrylamide gel. All Histags were C-terminally attached in this chapter so only the anti-His(C-term)-HRP antibody (R931-25, Invitrogen) was used.

2.8 Trypsin digestion and MALDI-TOF MS

The same protocol as described in previous chapter (Chapter 3 - 2.12) was applied. Identification was done with Mascot using the *E. coli* (strain K12) protein database, downloaded from Uniprot.

2.9 Fatty acid bioconversion assay

Hydroxylation of fatty acids (FAs) was assayed as follows. The reaction was performed in a volume of 1 ml, containing 100 mM potassium phosphate, pH 7.4, 0.1 mM oleic acid (from a 10 mM stock in DMSO), 1 U glucose-6-phosphate dehydrogenase (G6PDH), 5 mM glucose-6-phosphate (G6P) and 1 mM NADPH. The reaction was carried out for 20 min at 37 °C and 500 rpm in case of InaKtrBM3. For InaKtrCYPBMR (CYP indicating either CYPtr, CYPΔ32, CYPΔ25 or CYPnt), the reaction was incubated for 3 h at 37 °C and 500 rpm. The conversion was subsequently quenched with HCl and extraction was performed with and equal volume of diethyl ether (three times). The diethyl ether extracts were pooled, evaporated under vacuum and stored at – 20 °C until GC-MS analysis.

2.10 GC-MS analysis

Silylation and subsequent GC-MS analysis was performed as described in previous chapter (Chapter 3 - 2.14).

2.11 12-pNCA assay

After surface display, the cells were collected by centrifugation at 10 000 g and 4 °C for 10 min and washed three times with PBS. Thereafter, the cells displaying InaKtrBM3 were resuspended in PBS to a final OD_{600} of 50. OMVs were assayed after collection as described above. For measuring the activity of InaKtrBM3, the assay developed by Schwaneberg et al [364] was applied, with modifications. The reaction was carried out in a total volume of 250 µl, consisting of 100 mM Tris-HCl, pH 8.2 and either cells at a final OD_{600} of 2 or 25 µl resuspended OMVs. To this reaction, *p*-nitrophenoxydodecanoic acid (12-pNCA) (Molport) was added to a final concentration of 50 µM from a 5 mM stock solution in DMSO. After preheating at 30 °C, NADPH was added simultaneously to a final concentration of 0.1 mM. The formation of *p*-nitrophenol was monitored at 405 nm (ε = 13 200 $M^{-1}cm^{-1}$) and 30 °C for 30 min with shaking with the Bio-Rad Microplate Reader model 680.

2.12 Proteinase K accessibility assay

Cells were collected and washed with PBS after surface display as described in preparation of the 12-pNCA assay. To the resuspended cells, proteinase K was added to a final concentration of 0.1 mg/ml and incubated at 37 °C and 200 rpm for 1 h. Proteinase inhibitor was added and the cells were washed with 200 mM Tris-HCl, pH 8. Subsequently, the OM was collected as described by Hantke et al [365], with modifications. The washed cells were resuspended in 50 mM Tris-HCl, pH 8, containing 20 mM sucrose and 0.2 mM EDTA. To this suspension, lysozyme was added to a final concentration of 0.04 mg/ml. The suspension was incubated at room temperature for 10 min and protease inhibitor was added. To solubilize the IM, an equal volume of 50 mM Tris-HCl, pH 8, with 2 % Triton X-100 and 10 mM $MgCl_2$ was added. After incubation on ice for 30 min, cell debris was collected by centrifugation at 10 000 g and 4 °C for 10 min. The OM was pelleted from the supernatant by ultracentrifugation at 100 000 g and

4 °C for 1 h. The pellet was washed and after the second ultracentrifugation step, resuspended directly in Laemmli buffer.

In case of OMV display, proteinase K was added to resuspended OMV samples at a final concentration of 0.1 mg/ml. After incubation at 37 °C and 200 rpm for 1 h, an equal volume of 2X Laemmli was added for western blot analysis.

2.13 Nanoparticle Tracking Analysis

OMVs were quantified by means of nanoparticle tracking analysis (NTA), using the NanoSight LM10-HS microscope (Malvern Instruments Ltd), equipped with a 405 nm laser. The OMV suspension was diluted with PBS to fit within the measuring range of 3×10^8 and 1×10^9 particles/ml. For every sample, three movies of 30 s were recorded and the temperature was monitored. The recorded movies were processed with the NTA Analytical Software version 2.3. Three biological replicates were measured for assessing hypervesiculation. Subsequent statistical analysis was performed using the Graphpad Prism software.

2.14 OptiPrep™ density ultracentrifugation

The OMV pellet was resuspended in 2 ml 45 % OptiPrep™ in 10 mM Hepes, pH 7.4, 0.85 % (w/v) NaCl and transferred to the bottom of a 12 ml ultracentrifuge tube. Decreasing density layers were added, i.e. 2 ml 40 %, 2 ml 35 %, 2 ml 30 %, 2 ml 25 % and 1 ml 20 % OptiPrep™ in 10 mM Hepes, pH 7.4, 0.85 % (w/v) NaCl. Density ultracentrifugation was carried out at 141 000 g and 4 °C for 3 h. Samples of 1 ml were collected using the Gradient Station™ (BioComp) equipped with the Triax™ Flow Cell (BioComp). Next, these were diluted with 10 ml PBS and OMVs were pelleted by ultracentrifugation at 141 000 g and 4 °C for 3 h.

3 Results

3.1 The display of self-sufficient CYPs on the *E. coli* surface

Surface display was initiated by constructing the vector, containing the InaKtrBM3-encoding sequence for the surface display of CYP102A1 (from here on now denoted as BM3). Subsequently, the heme domain of BM3 was substituted by the gene encoding the truncated CYP52A13 as described in previous chapter, without the artificial N-terminal sequence. As the CYPtrBMR construct from previous chapter was not active, it was decided to make an alternative chimer, in analogy with the molecular lego approach described by Gilardi et al [273]. Instead of using the natural BM3 linker sequence, CYPtr was directly attached to BMR using a flexible linker of five Gly residues [274], [275]. We chose the BMR domain in the production host *E. coli* over the available CPRtr construct as the corresponding gene was predicted to be better suited for this recombinant host, based on the GenScript Rare Codon Analysis Tool and the online tool ATGme. The obtained plasmids were transformed in *E. coli* strain JM109, a strain often used for recombinant production of CYPs [229], [230]. After overnight expression at 28 °C, the cells were collected and lysate, inner membrane (IM) and outer membrane (OM) samples were isolated for western blot analysis (Figure 4.6). First, it is clear that both enzymes are produced, with InaKtrBM3 (152 kDa) present to a much higher level compared to InaKtrCYPtrBMR (154 kDa) in all fractions. Both constructs were present in the OM fraction, indicating that the surface display was successful for both enzymes. It must be noted that the used protocol for OM isolation rather leads to OM enrichment and might not result in completely pure OM so additional experiments are required to confirm surface display. In all lanes, multiple bands are seen for both enzymes. Higher MW bands indicate aggregation, which might be the result of extensive heating during the sample preparation (boiling of sample in Laemmli in preparation of SDS-PAGE) or due to contamination with formed inclusion bodies, although these should have been removed by the intermediate centrifugation step performed after sonication to remove cell debris. Multiple bands at a lower MW are seen as well, which shows that degradation occurs. Based on the MW of the lower bands, it was hypothesized that proteolysis of the linker sequences occurs. Proteolysis of the linker sequence in between InaKtr and BM3 or CYPtrBMR leads to a signal at 118 kDa or 120 kDa, respectively. Release of the BMR sequence by proteolysis of the linker in between the heme and reductase domain theoretically gives a signal around 65 kDa, as determined by the Expasy translate and protparam tools.

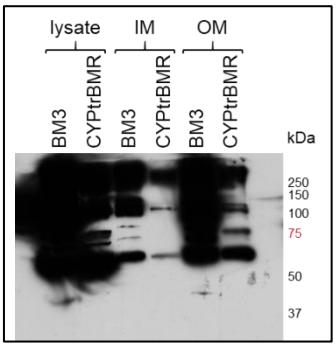

Figure 4.6: Western blot analysis of the collected cell fractions, that is the lysate, the inner membrane (IM) and the outer membrane (OM), after either InaKtrBM3 (152 kDa) or InaKtrCYPtrBMR (154 kDa) production in JM109.

Further characterization of the surface displayed enzymes was done by performing a whole-cell bioconversion assay. In case of InaKtrBM3, palmitic acid was used as the substrate. Following the bioconversion assay, the products were extracted and silylated for GC-MS analysis. The chromatogram obtained after a 20 min palmitic acid conversion by surface displayed InaKtrBM3 is shown in Figure 4.7. Two small peaks appeared in this chromatogram, i.e. at RT 7.971 min and 8.105 min. The MS spectra of both peaks are shown in Figure 4.8. Both spectra coincide with hydroxylated palmitic acid. The molecular ion of m/z 416 is not found, and it indeed is reported that the molecular ion of silylated OHFAs is only present in low abundance [366]. Instead the ion $[M-15]^+$, missing one methyl group, and the ions $[M-31]^+$ and $[M-105]^+$, resulting from interactions between the two functional groups, are seen. In the MS spectrum of Figure 4.8A, the ion with m/z 131 is highly abundant, indicating that palmitic acid is hydroxylated at the ω-2 position. The spectrum of Figure 4.8B shows a highly abundant ion with m/z 117, indicating hydroxylation at the ω-1 position. InaKtrBM3 thus actively converts palmitic acid mainly to 14-, and 15-hydroxypalmitic acid.

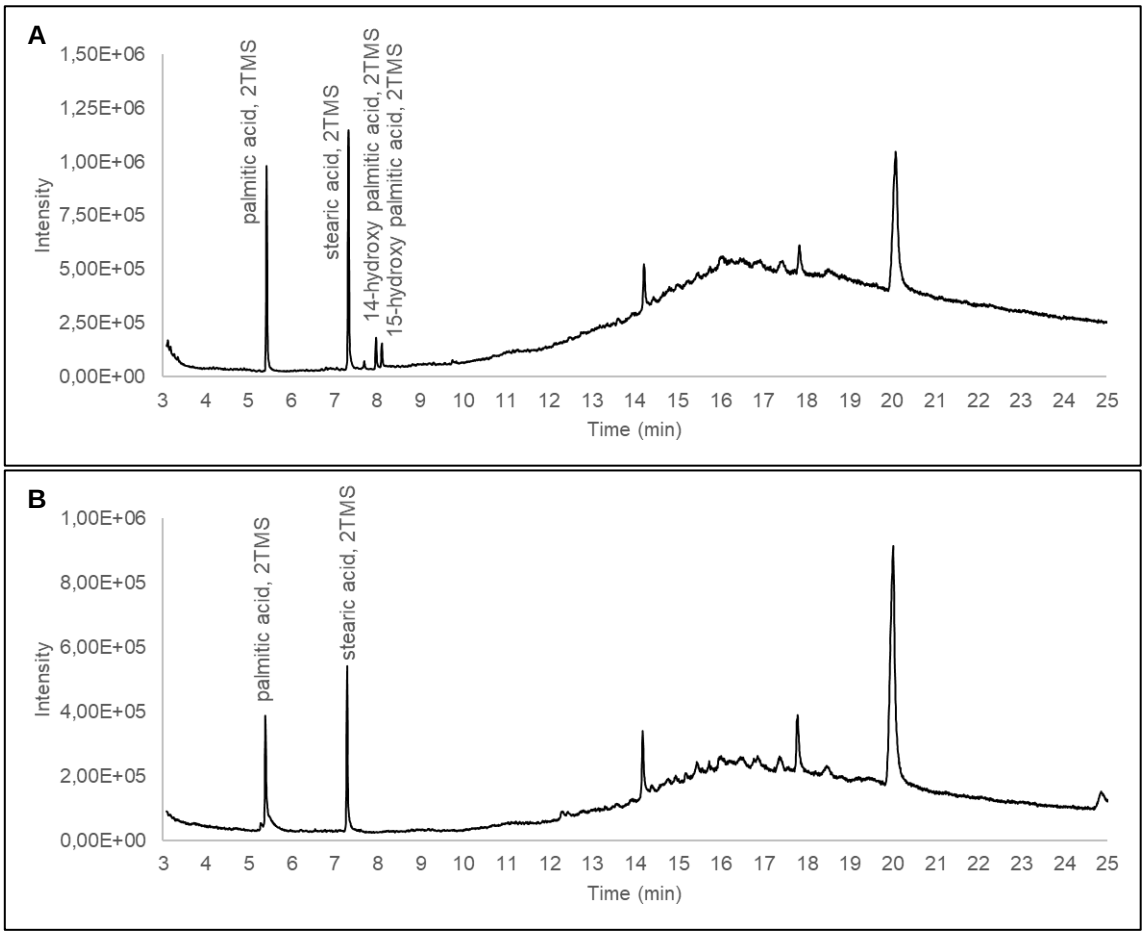

Figure 4.7: A: GC chromatogram of silylated extracts after a 20 min whole-cell palmitic acid bioconversion assay using JM109 cells, surface displaying InaKtrBM3. Palmitic acid was hydroxylated at its ω-1 and ω-2 position. B: GC chromatogram of silylated extracts of a blank, i.e. no palmitic acid substrate was added during bioconversion. The identified palmitic acid and stearic acid are contaminants, always observed by GC-MS analysis. Only peaks identified as silylated compounds, based on the presence in the MS spectrum of the characteristic peaks at m/z 73 and 75, are annotated.

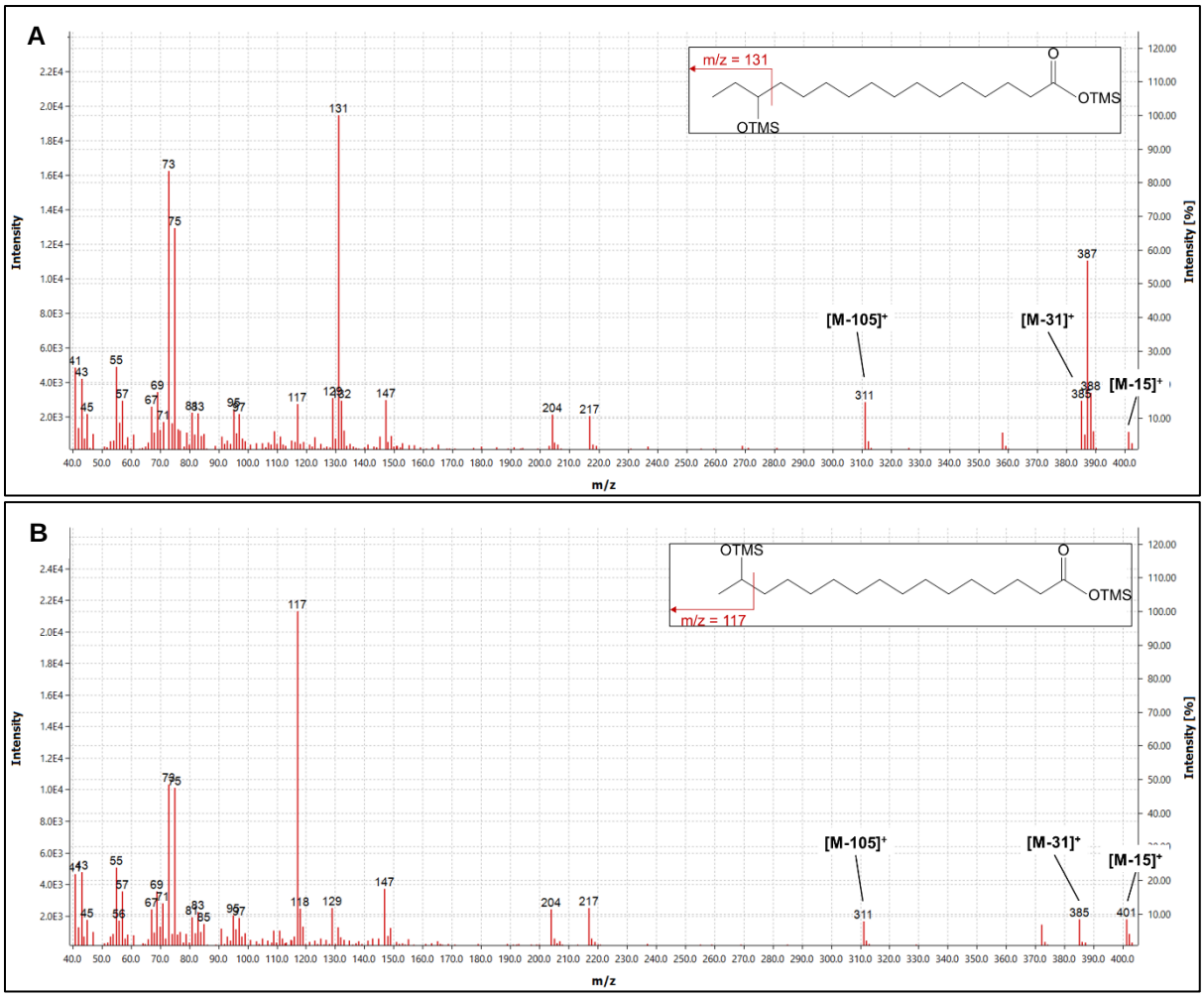

Figure 4.8: MS spectra of hydroxylated palmitic acid. A: MS spectrum of peak at RT 7.971 min (Figure 4.7), that is 14-hydroxypalmitic acid. B: MS spectrum of peak at RT 8.105 min (Figure 4.7), that is 15-hydroxypalmitic acid.

In case of InaKtrCYPtrBMR, oleic acid was used as the substrate for a whole-cell bioconversion assay. In one condition, a NADPH regeneration system was added. In the second condition, NADPH was discarded, serving as a blank. NADPH cannot migrate through the membrane so conversion linked to NADPH addition was hypothesized to be specific for the surface displayed construct. Unfortunately, no peaks coinciding with either OHFA or DCA were observed leading to the conclusion that no conversion occurred (data not shown).

CYPtr designed as described in previous chapter did not show any activity, neither in *P. pastoris,* nor in *E. coli*. Therefore, alternative truncations were performed, as described in Material and methods (see 2.3). In the same way as described for CYPtr, the heme domain of InaKtrBM3 was replaced by either CYPΔ32, CYPΔ25 or CYPnt. An expression test was performed and instead of isolating the lysate, IM and OM, surface display was assessed by means of a proteinase K accessibility assay. Only if the protein is present on the surface of *E. coli*, the protein will be degraded by proteinase K, as this protease cannot migrate through the membrane. Subsequent to proteinase K incubation, the OM was collected and analysed by SDS-PAGE (Figure 4.9A). A marked difference is seen between the OM samples isolated after proteinase K addition compared to a negative control (i.e. no addition of proteinase K). After

proteinase K incubation, all protein in the higher MW region seems to be absent. On the bottom of the Coomassie-stained gel, an intense band is observed. These bands were cut out from all eight lanes and analysed by MALDI-TOF MS. In all eight bands, the outer membrane porin (Omp) C was identified and in case of the untreated samples (i.e. no proteinase K was added), a mixture of both OmpC and OmpF was found, based on the peptide mass fingerprint. MS/MS analysis of the most abundant peptides resulted in the identification of OmpC alone in all cases. This confirms that indeed we have isolated the OM. The bands around 150 kDa were cut out as well for MALDI-TOF MS. Unfortunately, these could not be identified. Therefore, a western blot was executed (Figure 4.9B). Indeed all four constructs (InaKtrCYPtrBMR, InaKtrCYPΔ32BMR, InaKtrCYPΔ25BMR and InaKtrCYPntBMR) are present in the OM and were degraded upon proteinase K addition. All four chimers were thus successfully surface displayed. Of note, an alternative OM isolation protocol to the one used for OM isolation in Figure 4.6 was applied. This alternative protocol seems to lead to a much higher purity, as no contaminating signal is observed from the lysate or IM in the proteinase K-treated samples.

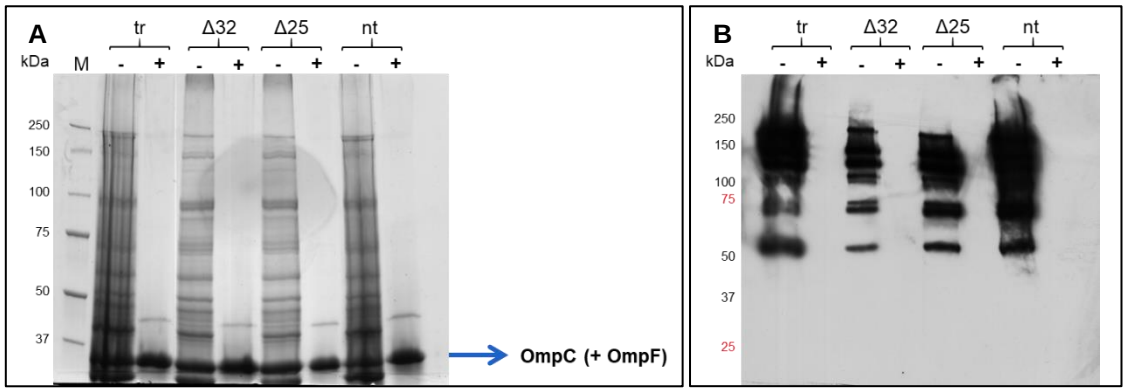

Figure 4.9: OM fraction, isolated after proteinase K accessibility assay. All four InaKtrCYPBMR, where CYP is either CYPtr, CYPΔ32, CYPΔ25 or CYPnt, were assayed. - : untreated. +: proteinase K-treated. A: SDS-PAGE analysis. The intense gel bands at the bottom of the Coomassie-stained gel were cut out and analysed by MALDI-TOF MS. Either OmpC or a mixture of OmpF and OmpC was identified. B: Western blot analysis.

A next step was to assay whether the surface displayed chimeric constructs were able to convert oleic acid to 18-hydroxy oleic acid and/or 1,18-octadecenedioic acid. As a blank, the NADPH regeneration system was left out of the reaction mixture. The resulting GC chromatograms from the CYPΔ32 construct are shown in Figure 4.10. The chromatograms from the other constructs were similar. Unfortunately, no OHFA nor DCA was observed so the alternative truncations did not lead to an active surface displayed chimer. Of note, oleic acid appeared to be consumed in absence of the NADPH regeneration system (Figure 4.10B). When only NADPH was discarded from the reaction mixture, thus still including G6P and G6PDH, oleic acid did not seem to be consumed in the blank (Figure 4.11). Oleic acid was thus degraded by β-oxidation in absence of the alternative energy source G6P. This highlights the disadvantage of working with whole-cells for the conversion of FAs, i.e. degradation of substrate (and formed products) by β-oxidation.

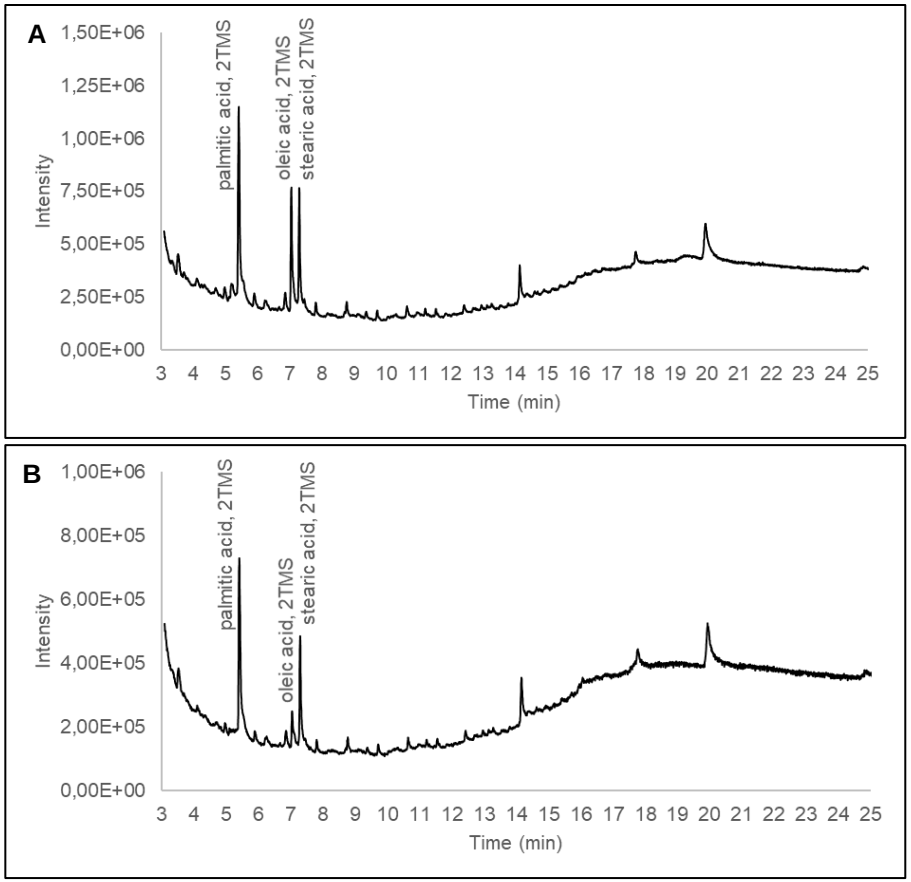

Figure 4.10: GC chromatogram of silylated extracts after whole-cell oleic acid bioconversion assay using JM109 cells, surface displaying InaKtrCYPΔ32BMR. A: Bioconversion assay including NADPH regeneration system. B: Bioconversion assay excluding NADPH regeneration system. No hydroxylated product of oleic acid was observed. Oleic acid appeared to be consumed in absence of an NADPH regeneration system. Only peaks identified as silylated compounds, based on the presence in the MS spectrum of the characteristic peaks at m/z 73 and 75, are annotated.

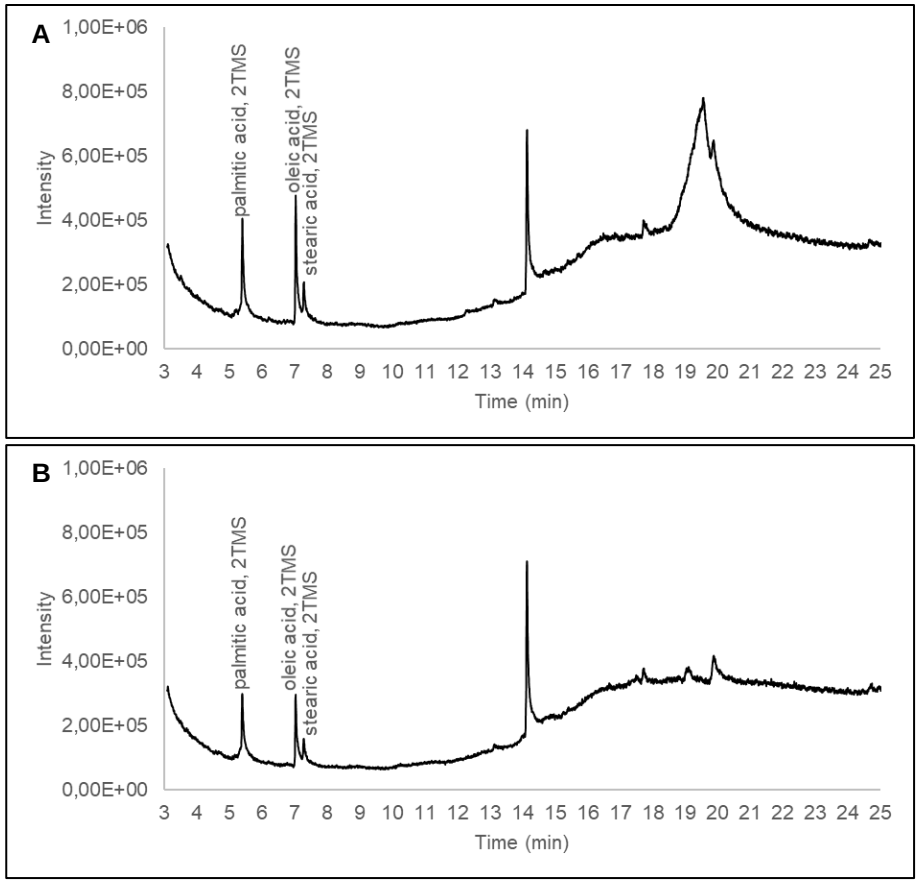

Figure 4.11: GC chromatogram of silylated extracts after whole-cell oleic acid bioconversion assay using JM109 cells, surface displaying InaKtrCYPΔ32BMR. A: Bioconversion assay including NADPH. B: Bioconversion assay excluding NADPH but including the regeneration components G6P and G6PDH. No hydroxylated product of oleic acid was observed. Only peaks identified as silylated compounds, based on the presence in the MS spectrum of the characteristic peaks at m/z 73 and 75, are annotated.

3.2 Taking it one step further: CYP display on outer membrane vesicles

3.2.1 Evaluation of hypervesiculating *E. coli*

Interest went to creating OMVs, displaying a self-sufficient CYP, in order to obtain a true *in vitro* system, with these OMVs serving as a nanoscale carrier. In normal conditions, *E. coli* produces only a low amount of OMVs. It is thus desired to induce hypervesiculation for an increased OMV yield [350]. Several mutations, as well as membrane intercalating molecules, have been reported to result in increased vesiculation. Single deletion mutants were ordered from the Keio collection and it was evaluated whether these indeed led to increased OMV formation. The single deletion mutant *tolA* was selected for the first OMV display experiments. OMVs were isolated when the cultures were in the early exponential phase and quantified by NTA (Figure 4.12). Both independent deletion mutants *tolA* 837 and *tolA* 838 were taken along in parallel and compared to the JM109 strain, used for surface display. At first sight, *tolA* 838 seemed to produce the highest amount of vesicles. However, this could not be statistically confirmed as the standard deviation was too big. In this experiment, two different types of ultracentrifuge tubes were used for OMV isolation, which may explain large variability. For further experiments, it was made sure that always the same type of ultracentrifuge tubes were used.

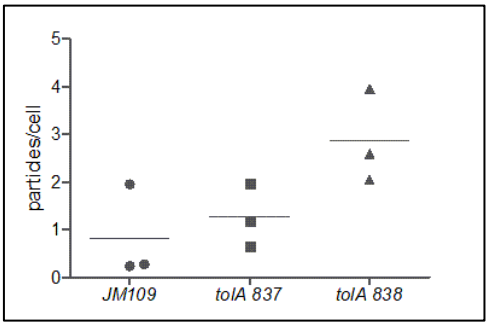

Figure 4.12: OMV quantification by NTA, displayed as particles per cell. The vesiculation did not vary significantly between the three different strains, as determined by one-way ANOVA (p = 0.0542).

The initial OMV display experiments using *tolA* 838 did not lead to great successes (see further, 0) so additional strains were ordered from the Keio collection, i.e. the two independent *nlpI* deletion mutants. Deletion mutants disrupting the σ^E stress response pathway were not ordered as neither a *degS* nor a *rseA* deletion mutant is available in this collection. The *degP* deletion mutant was reported to show impaired growth [367] and was therefore not considered. Of special interest was the hypervesiculating *E. coli* strain JC8031, where both *tolA* and *tolR* are knocked out. This strain showed a marked vesiculation increase. Additionally, this strain proved to be successful in the OMV display of recombinant proteins in several studies [349], [368], [369]. The same experiment as described for the *tolA* mutant strains was repeated, now including both *nlpI* deletion mutants (denoted at *nlpI* 863 and *nlpI* 864) and JC8031 (Figure 4.13A). In the early exponential phase, both *tolA* 838 and JC8031 produce a significantly higher amount of OMVs. Neither of the *nlpI* deletion mutants show hypervesiculation, this in contrast to previous reports [354], [355]. However, in these studies, OMVs were quantified after collection from a stationary phase culture. Schwechheimer et al hypothesized that NlpI suppresses PG endopeptidase activity. Deletion of *nlpI* leads to uncontrolled PG degradation,

to which the cell responds by an increased PG synthesis, even in the stationary phase, in order to survive. In the same study, it was observed that NlpI was specific for the PG endopeptidase Spr in the stationary phase, whereas the penicillin binding protein PBP4 was supressed in the exponential phase. Due to these differences between the exponential and stationary phase, OMVs from the different deletion mutants were isolated and quantified after overnight growth to the stationary phase (Figure 4.13B). Indeed, both *nlpI* deletion mutants now showed to produce an increased amount of OMVs, compared to JM109.

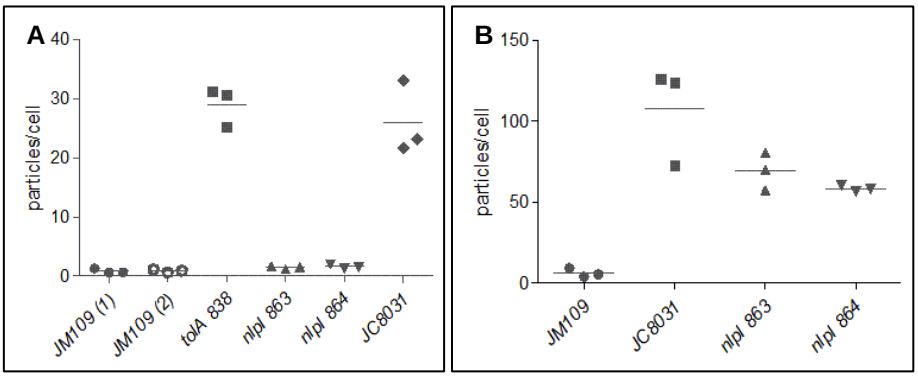

Figure 4.13: OMV quantification by NTA, displayed as particles per cell. A: OMVs were collected in the exponential phase. Three biological replicates were taken along in parallel. The vesiculation was significantly different between the different strains, as determined by one-way ANOVA ($p < 0.0001$). Tukey's multiple comparison test showed that both the *tolA* deletion mutant and JC8031 produced a significant higher amount of vesicles compared to all other strains, but did not differ significantly from each other. The three biological replicates at timepoint 1, that is JM109 (1), did not differ from the three biological replicates at timepoint 2, that is JM109 (2), proving reproducibility. B: OMVs were collected after overnight growth, that is in the stationary phase. Three biological replicates were taken along in parallel. The vesiculation was significantly different between the different strains, as determined by one-way ANOVA ($p = 0.0004$). All three deletion mutants showed a significantly higher vesicle production compared to JM109, based on Tukey's multiple comparison test. JC8031 showed to vesiculate to a higher extend than *nlpI* 864, but not *nlpI* 863, based on Tukey's multiple comparison test. *nlpI* 863 and 864 vesiculation did not significantly vary.

Several antibiotics have been shown to induce OMV production. It was investigated whether the JM109 strain showed increased OMV production adding polymyxin B. The subinhibitory concentration of 0.75 µg/ml, as applied by Manning and Kuehn [357], was added to an exponential phase culture of JM109. After an incubation time of 3 h, the medium was collected for OMV isolation and quantification by NTA (Figure 4.14). Indeed no growth inhibition was observed using this polymyxin B concentration, as the OD_{600} did not differ after a 3 h incubation (Figure 4.14A). Growth inhibition was observed only from a concentration of 1.5 µg/ml on (data not shown). Upon incubation with polymyxin B, a significant increased amount of OMVs was observed (Figure 4.14B). However, when compared to JC8031, this increased vesiculation appeared to be negligible (Figure 4.14C). Indeed, it has been reported that factors such as antibiotics and PQS do not increase the vesiculation to yields, useful for pharmaceutical and biotechnological applications [350].

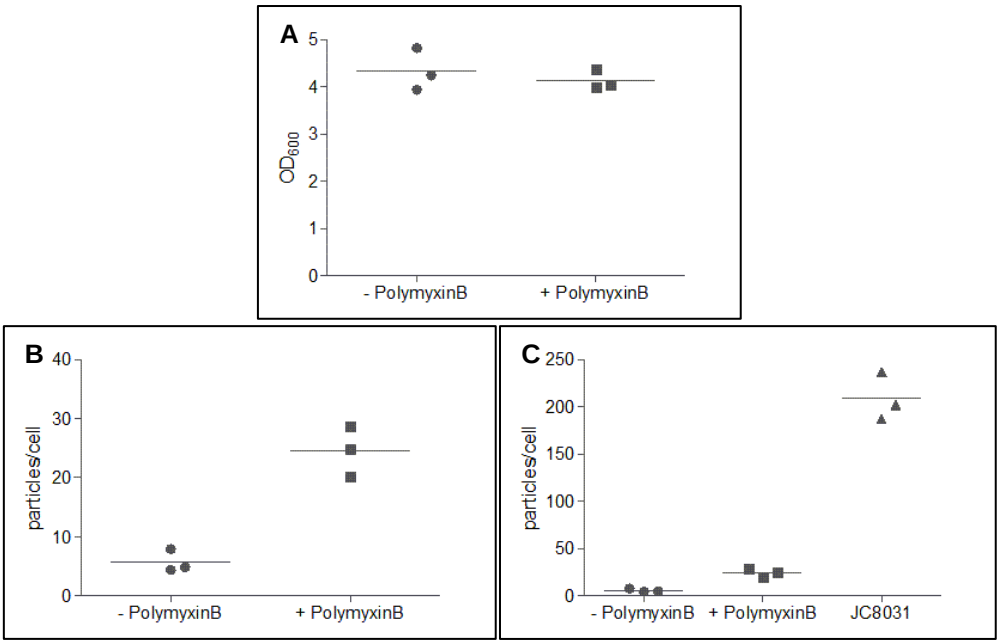

Figure 4.14: Polymyxin B induced vesiculation. A: Optical density of JM109 after 3 h growth post polymyxin B addition to one of two cultures, executed in triplicate. The optical density did not differ significantly in absence or in presence of 0.75 µg/ml polymyxin B, determined by an unpaired t-test ($p = 0.5262$). B: OMV quantification by NTA, displayed as particles per cell, after induction of OMV biogenesis by polymyxin B addition, compared to a non-induced control. Polymyxin B induction significantly increased vesiculation, determined by an unpaired t-test ($p = 0.0022$). C: OMV quantification by NTA, displayed as particles per cell. OMVs were collected after growing the JC8031 strain for the same period of time as the polymyxin B induced cultures (performed in triplicate). The vesiculation increase by polymyxin B addition became insignificant when including the hypervesiculating strain, determined by one-way ANOVA ($p < 0.0001$), followed by Tukey's multiple comparison test.

3.2.2 Proof-of-concept: the display of CYP102A1 on the surface of OMVs

In order to investigate whether OMV display of an active self-sufficient CYP is feasible, it was decided to first start with expression tests using the InaKtrBM3 construct, as surface display of this enzyme was already accomplished and was confirmed in this dissertation. When starting these OMV display experiments, only the *tolA* deletion mutants were available and thus expression tests were initially performed with these strains. A first 50 ml expression test was performed where *tolA* 837 and 838 were grown in parallel. InaKtrBM3 expression was induced at an OD_{600} of 0.5 with 1 mM IPTG and the cells were grown further to the late exponential phase, at which point the OMVs were isolated and analysed by western blot (Figure 4.15A). No signal was observed, so no OMVs displaying InaKtrBM3 were isolated. Continuing with *tolA* 838, some alternative expression conditions were investigated. Additionally, in all conditions, a larger culture was grown in order to increase the amount of isolated OMVs. In a first setup, OMVs were isolated after an expression phase of 3 h. Secondly, an overnight expression phase was performed. Thirdly, instead of IPTG induction, cells were grown in autoinduction medium, as it has been stated to lead to increased densities [360], [370]. Lastly, overnight expression at 20 °C instead of 28 °C was performed in order to assess whether this would lead to improved yield of properly folded enzyme and thus increase the yield displayed on OMVs. Unfortunately, none of these conditions led to successful isolation of OMVs displaying InaKtrBM3, as no signal was observed by western blot analysis (Figure 4.15B). To make sure that the enzyme was effectively produced and surface displayed, the OM was isolated and analysed in parallel. In all cases where expression was induced with IPTG, InaKtrBM3 was present around the right MW, together with several degradation bands. It was already mentioned before that the linker sequences seem to be prone to proteolysis. In case of autoinduction, only one band is seen and this at a lower MW. Autoinduction appears thus to be unsuitable for surface display of this enzyme.

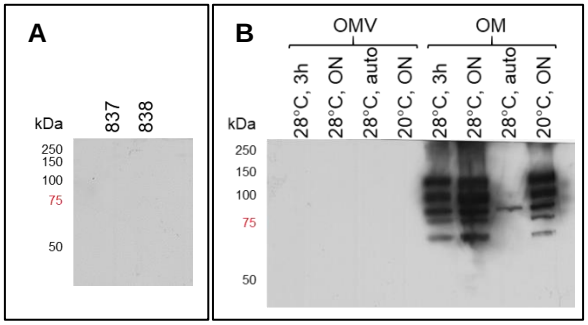

Figure 4.15: Western blot analysis after InaKtrBM3 (152 kDa) production in the *tolA* deletion mutant(s). A: OMV samples isolated after expression in both *tolA* 837 and *tolA* 838. B: OMV and OM samples after InaKtrBM3 production in *tolA* 838, applying four different expression conditions.

No success was booked using the *tolA* deletion mutant so other strains were selected, as described above. The *tolRA* double deletion mutant JC8031 produced the highest amount of OMVs per cell, apart from *tolA* 838, so this strain was subsequently used for InaKtrBM3 production. Of note, from here on now, TB medium instead of LB medium was used when collecting OMVs of an exponential phase culture, as it was observed that higher cell densities were obtained in TB medium. A first expression test showed that InaKtrBM3 was produced and successfully surface displayed after a 3 h expression phase. Instead of performing a

conversion assay, followed by GC-MS analysis, a quick and easy continuous spectrophotometric enzymatic assay was used for assessing activity. Schwaneberg et al developed an enzymatic assay, specifically for BM3 [364]. The assay is based on the conversion of *p*-nitrophenoxycarboxylic acids (pNCAs) by ω-hydroxylation, thereby releasing ω-oxycarboxylic acids and the chromophore p-nitrophenol. Assaying intact JC8031 cells after InaKtrBM3 production, the cells proved to produce p-nitrophenol (Figure 4.16A). As a blank, NADPH was left out of the reaction mixture. NADPH is not able to traverse the membrane so observed activity upon NADPH addition must be the result of surface displayed InaKtrBM3. The OMVs were isolated and assayed as well, showing 12-pNCA conversion activity (Figure 4.16B). An extra wash step was performed in order to remove any residual medium and these washed OMVs were again assayed (Figure 4.16C). Unfortunately, no activity was observed anymore. The OMVs were further analysed by SDS-PAGE and compared to OMVs of JC8031, not producing any recombinant protein. In parallel, the OM was isolated and taken along in the SDS-PAGE analysis (Figure 4.17A). In the OM lanes, a band appears around 150 kDa, which is absent in case of JC8031, not producing recombinant protein. This band was not seen in the OMV lanes. The band was cut out and by MALDI-TOF MS it was confirmed to be InaKtrBM3. An expression test producing InaKtrBM3 overnight in LB medium, in analogy to the surface display experiments, was performed as well, showing similar results. Due to the fact that activity was lost upon extra washing of the OMV pellet, a medium sample was analysed by SDS-PAGE (Figure 4.17B). Here, a band appeared at a slightly lower MW. Again, this band was cut out and analysed by MALDI-TOF MS. It was identified as BM3 and no InaKtr-specific peptides could be found in the MS spectrum (see Addendum). Taken together with the aforementioned activity in the crude OMV sample, it was clear that a significant amount of active BM3 is cleaved of the InaKtr anchor and released in the medium.

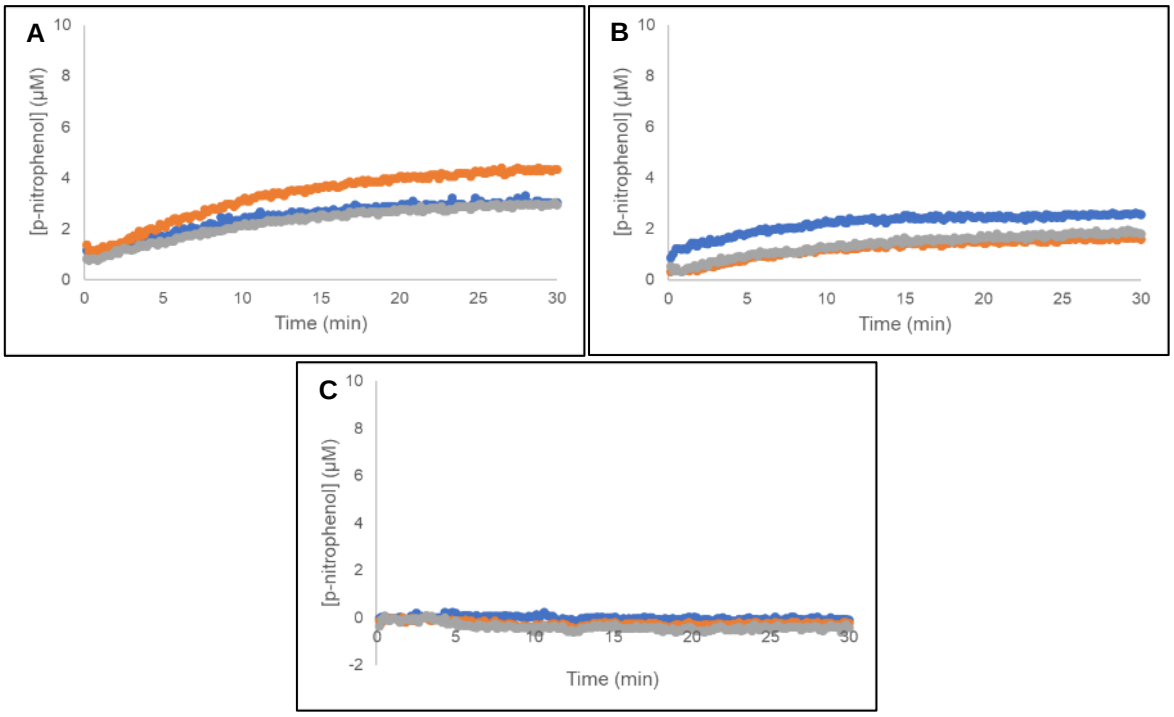

Figure 4.16: 12-pNCA assay after InaKtrBM3 production in JC8031 for 3 h in TB medium. A: cells. B: crude OMVs. C: washed OMVs.

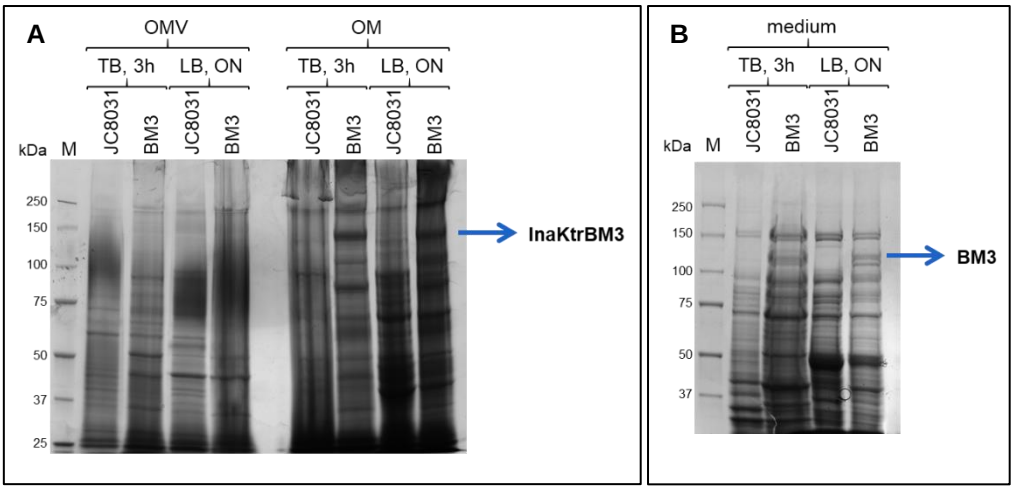

Figure 4.17: SDS-PAGE analysis after InaKtrBM3 production in JC8031 for 3 h in TB medium or overnight in LB medium. A JC8031 culture not producing recombinant protein was taken along in parallel. A: OMV and OM samples The band appearing around 150 kDa was identified as InaKtrBM3 by MALDI-TOF MS. B: Concentrated medium samples. The band appearing in between 100 kDa and 150 kDa was identified as BM3 and no peptides, specific for the InaKtr anchor were found (see Addendum).

During these experiments, it was evident that the surface display of InaKtrBM3 in JC8031 led to growth impairment as the optical density was significantly lower than the density reached by JC8031, not producing recombinant protein, in the same period of time. An OD_{600} of 2.15 was reached after a 3 h expression phase, compared to 5.16 for JC8031, not producing recombinant protein. After overnight expression, an OD_{600} of 2.72 was measured, compared to 6.12 for JC8031, not producing recombinant protein. This slower growth was already seen during the growth phase, indicating leaky expression. Efforts were done to resolve these issues. On the one hand, glucose was added and IPTG concentration for induction was reduced to 0.3 mM. On the other hand, *inaKtrBM3* was cloned into the vector pBAD/Myc-HisA, where the gene is under the control of the more tightly regulated arabinose-induced *araBAD* promotor. In both cases, optical densities were reached comparable to JC8031, with the construct under the control of the *araBAD* promotor showing superior growth over the IPTG-induced culture. Unfortunately, no OMVs were found to contain InaKtrBM3, neither after a 3 h expression phase, upon collection in the late exponential phase and after overnight expression (Figure 4.18A, B and C, respectively). From these experiments, it was confirmed that a significant amount of BM3 was cleaved of the InaKtr anchor and released in the medium. A concentrated sample of the OMV supernatant showed the most intense band at a lower MW than the expected MW of 152 kDa upon collection at late exponential phase (Figure 4.18B). Also after overnight expression, a lower MW band is observed in the supernatant (and in the crude OMV sample collected after expression of InaKtrBM3_pMAL) (Figure 4.18C).

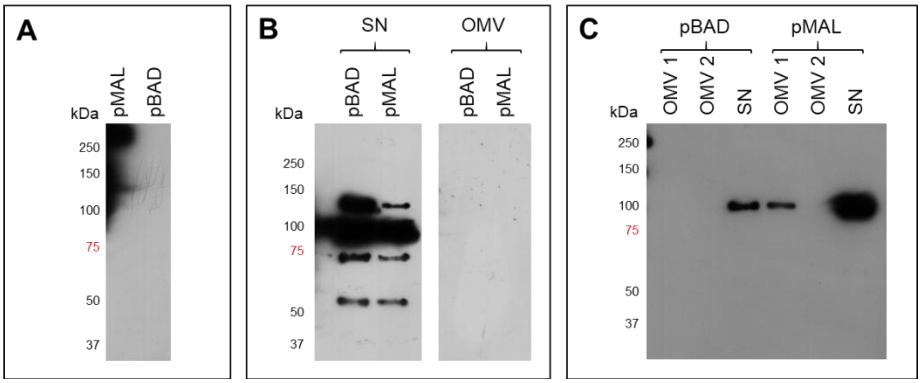

Figure 4.18: Western blot analysis of OMV samples derived from JC8031 after producing InaKtrBM3 (152 kDa), either under the control of the *tac* promotor (pMAL) or under the control of the *araBAD* promotor (pBAD). A: 3 h expression phase. B: OMV collection in the late exponential phase. A concentrated sample of the OMV supernatant (SN) was taken along in parallel. C: Overnight expression phase. An OMV sample was taken before and after an extra wash step (i.e. OMV 1 and OMV 2, respectively), together with a concentrated sample of the OMV supernatant.

No spontaneously formed OMVs displaying InaKtrBM3 were isolated so far. Therefore, another path was taken, inspired by the application of OMVs as a vaccine. In the development of OMV-based vaccines, not only spontaneously formed OMVs are investigated (from here on now referred to as sOMV). Rather, OMVs are extracted after cell collection using either detergents, resulting in so-called dOMVs, or detergent-free extraction methods [348]. These detergent-free extraction methods are reported to result in more native-like OMVs (from here on now referred to as extracted OMVs or eOMVs). Van de Waterbeemd et al reported the use of EDTA, a chelating agent that destabilizes the OM and thereby releasing OMVs, in combination with a genetic modification resulting in a loosely attached OM. This approach resulted in a good eOMV yield, and this yield was superior to the sOMV yield [363]. Therefore, EDTA extraction of eOMVs from the JC8031 strain, recombinantly producing InaKtrBM3, was performed. Both the pMAL and the pBAD constructs were taken along in parallel, as well as the original JM109 strain, surface displaying InaKtrBM3. After isolation of the eOMVs by ultracentrifugation, followed by a wash step, the eOMVs were assayed for activity (Figure 4.19). Expression of *inaKtrBM3* under the control of the *tac* promotor (pMAL) led to eOMVs showing p-nitrophenol formation to equilibration (Figure 4.19A). A moderate activity was seen in case of arabinose induction (pBAD) (Figure 4.19B). No activity was seen when producing InaKtrBM3 in the JM109 strain (Figure 4.19C). Indeed, western blot analysis shows that InaKtrBM3 is present in both JC8031-derived eOMVs, whereas the enzyme is absent from eOMVs, derived from JM109 (Figure 4.19D).

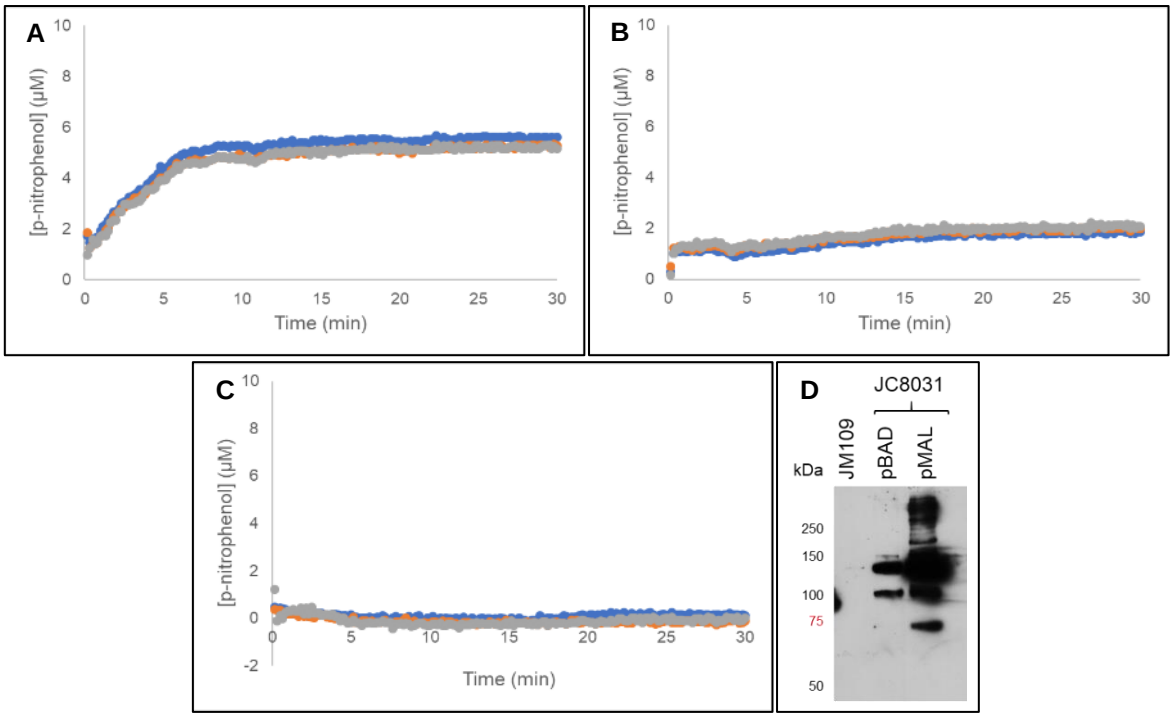

Figure 4.19: eOMVs extracted after InaKtrBM3 (152 kDa) production. A: 12-pNCA assay. JC8031-derived eOMVs. Expression was under the control of the *tac* promotor (pMAL). B: 12-pNCA assay. JC8031-derived eOMVs. Expression was under the control of the *araBAD* promotor (pBAD). C: 12-pNCA assay. JM109-derived eOMVs. D: Western blot analysis.

Although success was booked with eOMVs, interest still went to obtaining sOMVs, as collection of these vesicles does not require an extra extraction step. To this end, both InaKtrBM3_pMAL and InaKtrBM3_pBAD were transformed in the *nlpI* deletion mutant (*nlpI* 863). After overnight expression, allowing the cells to grow to the stationary phase (as only then, hypervesiculation was observed), both sOMVs and eOMVs were collected. sOMVs were collected from the JM109 strain as well, where hypervesiculation was induced by polymyxin B addition. eOMV extraction from the JM109 strain was again performed, but now after overnight expression. The JC8031 strain was taken along in parallel. Next to the 12-pNCA assay, OMV samples were analysed by western blot (Figure 4.20A). To verify if the InaKtrBM3 was truly displayed on the OMV surface, a proteinase K accessibility assay was performed, after which again a western blot analysis was executed (Figure 4.20B). These analyses showed that only eOMVs contained InaKtrBM3 and that these were surface displayed, as proteinase K completely degraded the enzyme. Neither using the *nlpI* deletion mutant, nor OMV induction by polymyxin B addition resulted in successfully obtaining sOMVs. The JC8031 strain also resulted in the best eOMV yield, especially when *inaKtrBM3* was under the control of the *tac* promotor. Some minor signal was also seen in the eOMV sample extracted from *nlpI* 863 using the pMAL construct. The results seen on western blot were further confirmed by a 12-pNCA assay. In case of JC8031, the eOMVs were clearly active (Figure 4.21A and B), with the eOMVs after IPTG induction reaching equilibrium very quickly (Figure 4.21A). The sOMVs derived from JC8031 seemed to show minor activity (Figure 4.21C and D), although no signal was seen by western blot. OMVs derived from the JM109 strain did not contain any InaKtrBM3 according to the western blot, and indeed the sOMVs did not show any activity (Figure 4.22B). However, some minor activity was seen in the eOMV sample (Figure 4.22A). Based on the

western blot analysis, InaKtrBM3 was present in the eOMVs derived from *nlpI* 863 after IPTG induction. This was confirmed by the 12-pNCA assay, as a small amount of p-nitrophenol was formed (Figure 4.23A). Also the eOMVs after arabinose induction seem to be slightly active (Figure 4.23B). Neither of the sOMV samples showed any p-nitrophenol formation. These experiments show that the extraction of OMVs using EDTA, in combination with mutations impairing envelope stability, result in the highest OMV yields.

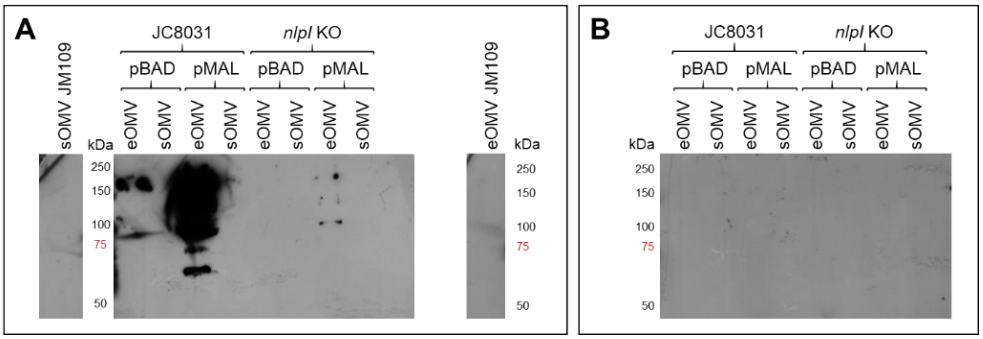

Figure 4.20: A: Western blot analysis of eOMVs and sOMVs, either derived from JM109, JC8031 or *nlpI* deletion mutant 863. In the last two strains, two different constructs (pBAD, i.e. *araBAD* promotor, or pMAL, i.e. *tac* promotor) were taken along. B: Western blot analysis of the proteinase K treated OMV samples.

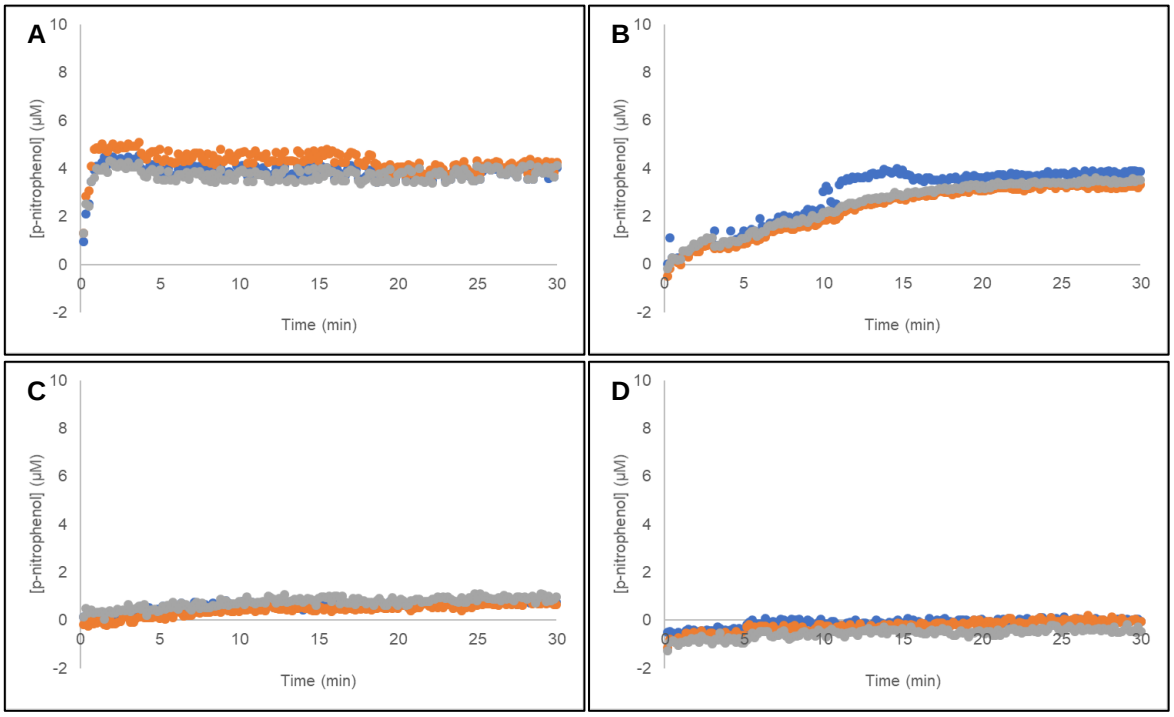

Figure 4.21: 12-pNCA assay using OMVs isolated after InaKtrBM3 production in JC8031. A: eOMVs. Expression was under the control of the *tac* promotor. B: eOMVs. Expression was under the control of the *araBAD* promotor. C: sOMVs. Expression was under the control of the *tac* promotor. D: sOMVs. Expression was under the control of the *araBAD* promotor.

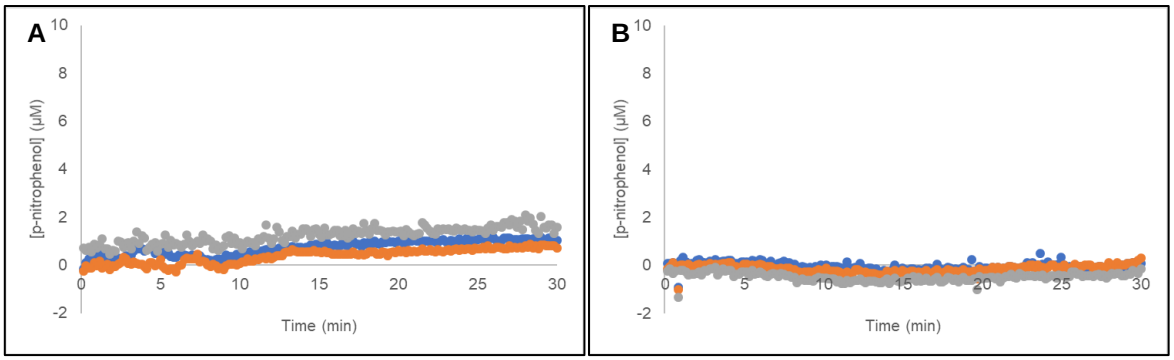

Figure 4.22: 12-pNCA assay using OMVs isolated after InaKtrBM3 production in JM109. A: eOMVs extracted after overnight expression. B: sOMVs collected 6 h post polymyxin B induction.

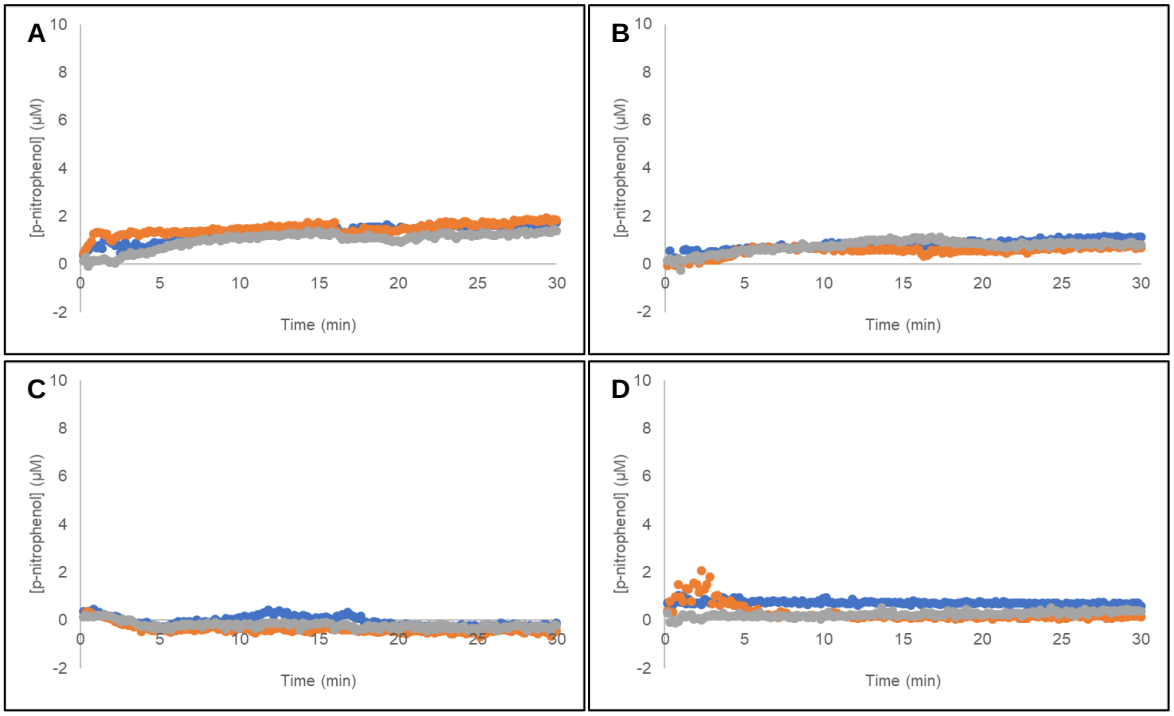

Figure 4.23: 12-pNCA assay using OMVs isolated after InaKtrBM3 production in *nlpI* deletion mutant 863. A: eOMVs. Expression was under the control of the *tac* promotor. B: eOMVs. Expression was under the control of the *araBAD* promotor. C: sOMVs. Expression was under the control of the *tac* promotor. D: sOMVs. Expression was under the control of the *araBAD* promotor.

Although no signal coinciding with InaKtrBM3 was observed by western blot, the seemingly small activity observed in the 12-pNCA assay prompted us to further pursue the sOMV approach. So far, the culture supernatant was filtered using a 0.22 µm filter before sOMV collection by ultracentrifugation, as this was an established protocol within the lab. This pore size is smaller than the upper limit of the OMV size. Indeed, scanning literature learned that 0.45 µm filters are often used instead of 0.22 µm filters [351], [371]. We wondered whether the use of 0.45 µm filters could increase the OMV yield, and thereby obtaining a sufficient amount of active OMVs. A culture was set up where InaKtrBM3 was produced in JC8031 upon IPTG induction. The culture supernatant was divided in two. One half was filtered using a 0.22 µm filter, whereas the other half was filtered using a 0.45 µm filter. In this way, it was made sure that no differences could be ascribed to biological variations between different cultures. The isolated sOMVs were subsequently assayed for 12-pNCA hydroxylation and analysed by western blot (Figure 4.24A and B, respectively). An increasing absorbance was observed after

use of the 0.45 µm filter, whereas no such increase was seen in case of the 0.22 µm filter. Western blot analysis showed a signal in the sOMV sample collected after using the 0.45 µm filter and not in case of the 0.22 µm filter. However, this signal was observed at a lower than expected MW of 152 kDa. No residual medium was present as OMVs were washed as mentioned before. This western blot analysis was not performed immediately after sOMV collection so the enzyme might have degraded during storage. Both sOMV samples were further analysed by NTA (Figure 4.24C). Indeed, relatively more particles with sizes around 200 nm were recovered. The overall particle concentration is not comparable to one another due to differences in the OMV resuspension volume.

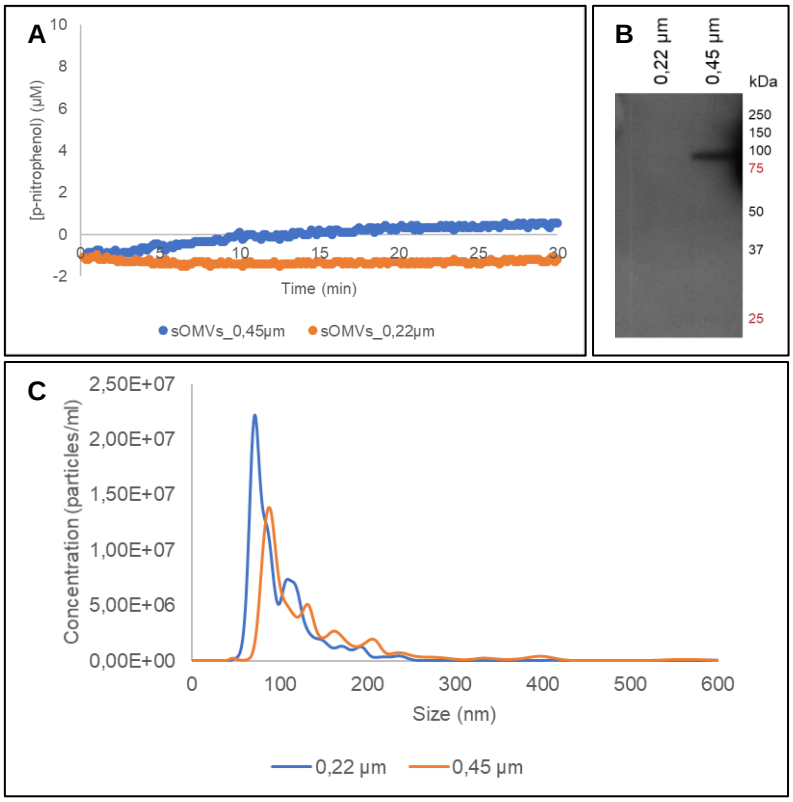

Figure 4.24: sOMVs isolated after InaKtrBM3 expression in JC8031 under the control of the *tac* promotor. One half of the medium was filtered with a 0.22 µm filter. The other half was filtered with 0.45 µm filter. A: 12-pNCA assay. B: Western blot analysis. C: NTA analysis.

As InaKtrBM3 seemed to be present in sOMVs, collected after filtering the supernatant with a 0.45 µm filter, all strains and constructs, mentioned above, were reinvestigated. Both eOMVs and sOMVs were collected, now filtering the respective supernatant with a 0.45 µm filter, and tested for activity by the 12-pNCA assay. First, the JC8031 strain was revisited in order to confirm that the sOMVs, obtained after 0.45 µm filtration indeed were active and the arabinose-induced construct was taken along in parallel (Figure 4.25). In Figure 4.25A, p-nitrophenol formation was clearly seen, thereby confirming that active sOMVs were successfully obtained. It must be noted that the medium appeared to be somewhat viscous and eOMV extraction was no longer possible as cells lysed upon incubation with the EDTA-rich buffer. It was noticed that every time the JC8031 strain with the construct under the control of the *tac* promotor was grown, starting from a glycerol stock made from the preceding preculture, a slightly lower OD_{600} value was reached upon collection in late exponential phase. This raised the question whether these OMVs truly are sOMVs resulting from OM blebbing, or are these the result of (partial) cell lysis. There was no clear activity seen for any other sOMV sample (Figure 4.26B and D). The JC8031-derived eOMVs collected after InaKtrBM3 production, under the control of *araBAD* were active (Figure 4.25B), as well as the *nlpI* 863-derived eOMVs collected after InaKtrBM3 production under the control of the *tac* promotor (Figure 4.26A), which was also seen after filtration with a pore size of 0.22 µm. Now, also the *nlpI* 863-derived eOMVs under the control of *araBAD* showed activity (Figure 4.26C). Neither the vesicles derived from *tolA* 838 nor derived from the polymyxin B-induced JM109 strain actively converted 12-pNCA (data not shown).

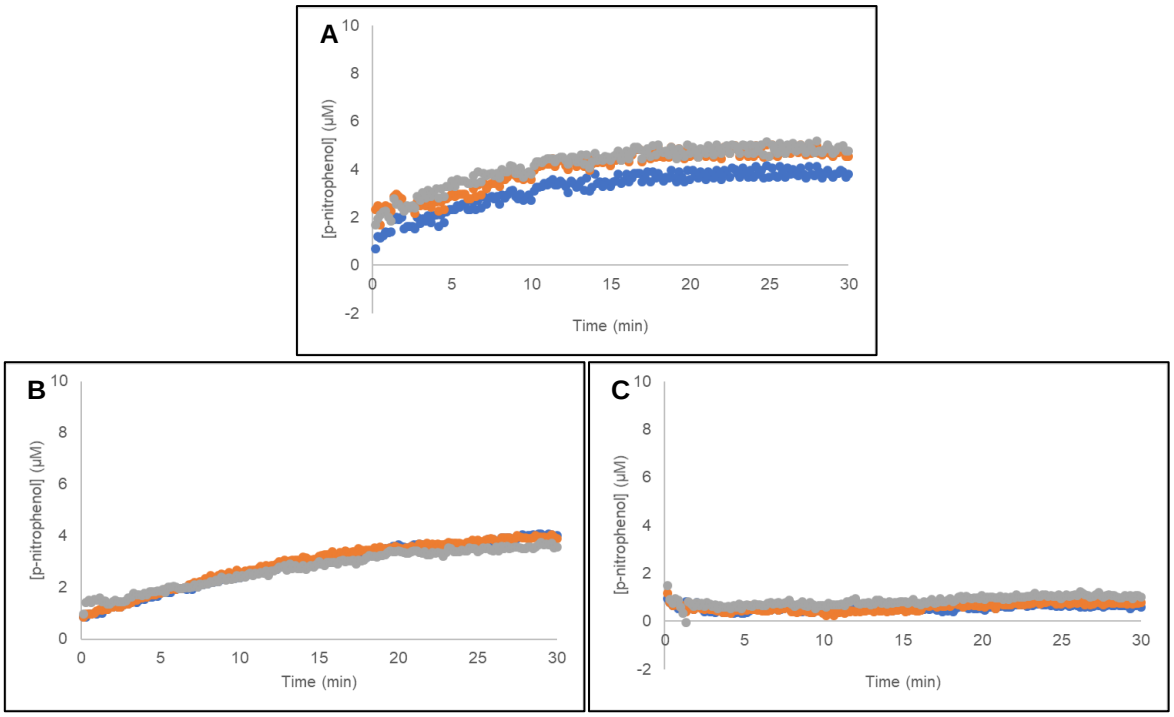

Figure 4.25: 12-pNCA assay using OMVs isolated after InaKtrBM3 expression in JC8031. Filtration was done with a 0.45 µm filter instead of a 0.22 µm filter. A: sOMVs. Expression was under the control of the *tac* promotor. B: eOMVs. Expression was under the control of the *araBAD* promotor. C:sOMVs. Expression was under the control of the *araBAD* promotor.

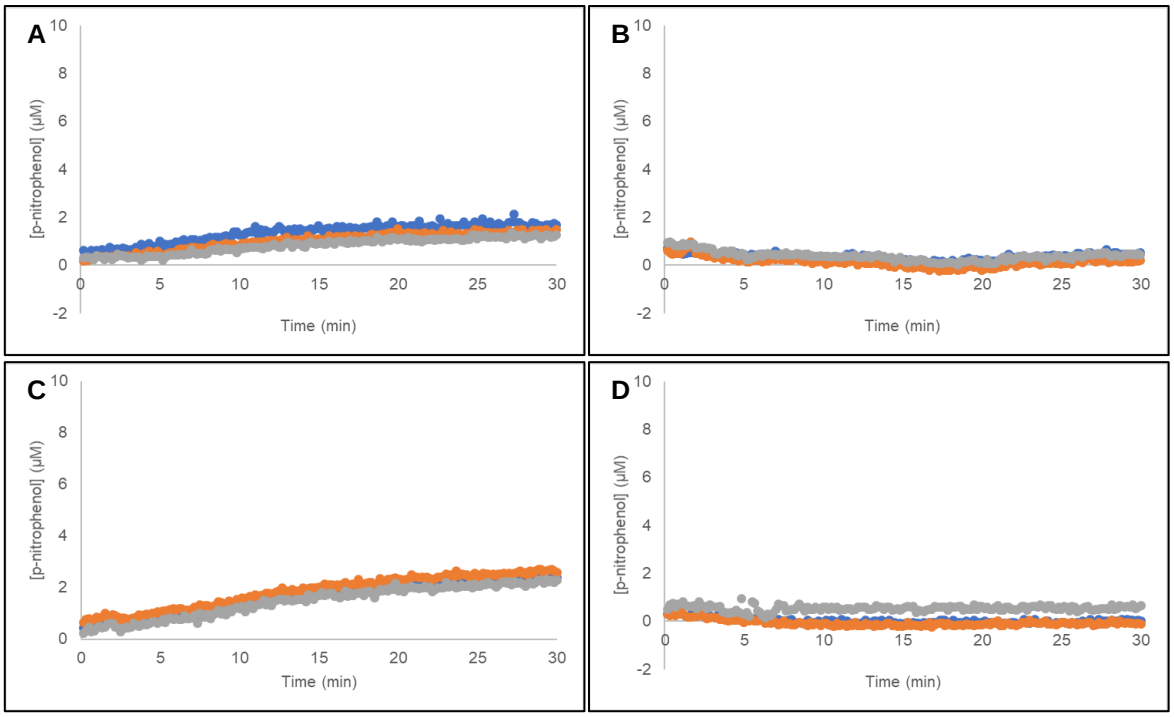

Figure 4.26: 12-pNCA assay using OMVs isolated after InaKtrBM3 production in *nlpI* deletion mutant 863. Filtration was done with a 0.45 µm filter instead of a 0.22 µm filter. A: eOMVs. Expression was under the control of the *tac* promotor. B: sOMVs. Expression was under the control of the *tac* promotor. C: eOMVs. Expression was under the control of the *araBAD* promotor. D: sOMVs. Expression was under the control of the *araBAD* promotor.

To further characterize the isolated OMVs, purification by OptiPrep™ density ultracentrifugation was attempted. The sOMVs were isolated as before from the JC8031 strain producing InaKtrBM3 upon IPTG induction and were subsequently loaded on the bottom of an OptiPrep™ gradient. Ultracentrifugation was performed for 3 h, after which 1 ml fractions were collected automatically, while measuring the absorbance at 260 nm (Figure 4.27A). This graph shows two spikes of absorbance increase, that is in fraction 6 and 8. In this graph, a stepwise absorbance increase was seen, indicating that no linear gradient was obtained. sOMVs (or other types of membrane structures) were collected from each fraction and analysed by western blot (Figure 4.27B). Of note, the bottom layer was not collected by the fraction collector, i.e. fraction 11, and appeared to be very viscous upon Laemmli buffer addition whereby analysis by SDS-PAGE and western blot was not possible. Based on the western blot analysis, the higher density fractions 8, 9 and 10 contained InaKtrBM3. However, vesicles are expected to float to lower density fractions. The InaKtrBM3-containing fractions were further investigated by SDS-PAGE (Figure 4.27C). In all three fractions, intense bands are present around the expected MW of OmpF, OmpC and OmpA, which have been used as markers for OMVs due to their large abundance in the OM and OMVs. The gel bands were cut out and analysed by MALDI-TOF MS. Only OmpC was found in the upper band and not a mixture of OmpF and OmpC. The lower band was identified as OmpA. As these Omp proteins are also present in these higher density fractions, together with the fact that no linear gradient was obtained, it was concluded that a longer centrifugation time is required for proper vesicle purification. The same protocol was used for the purification of eOMVs, derived from JC8031 after InaKtrBM3 production,

induced by arabinose (Figure 4.28). Similarly, the higher density fractions showed to be positive for InaKtrBM3 and OmpC and OmpA were found in the same fractions.

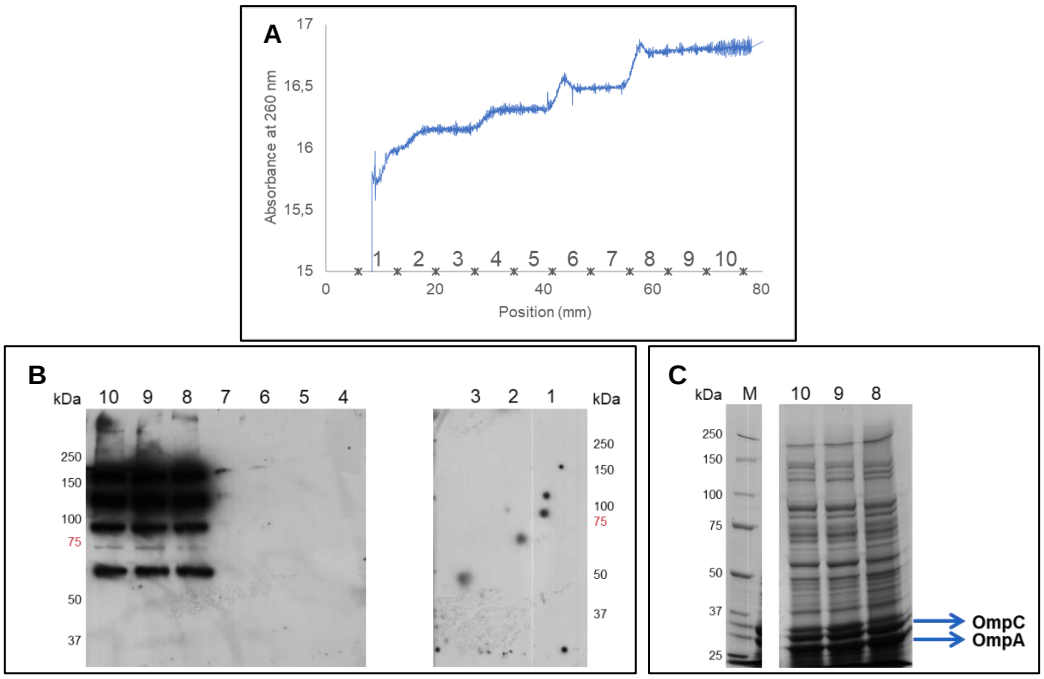

Figure 4.27: OptiPrep™ density ultracentrifugation sOMVs, derived from JC8031 after InaKtrBM3 production under the control of the *tac* promotor. A: Absorbance at 260 nm of OptiPrep™ gradient after centrifugation for 3 h at 141 000 g. 1 ml fractions were collected from top to bottom and were numbered accordingly. B: Western blot analysis of the collected fractions after OptiPrep™ density ultracentrifugation. C: SDS-PAGE analysis of the InaKtrBM3-positive fractions. The intense gel bands in between 25 kDa and 37 kDa were cut out and analysed by MALDI-TOF MS. The upper band was identified as OmpC, the lower band as OmpA.

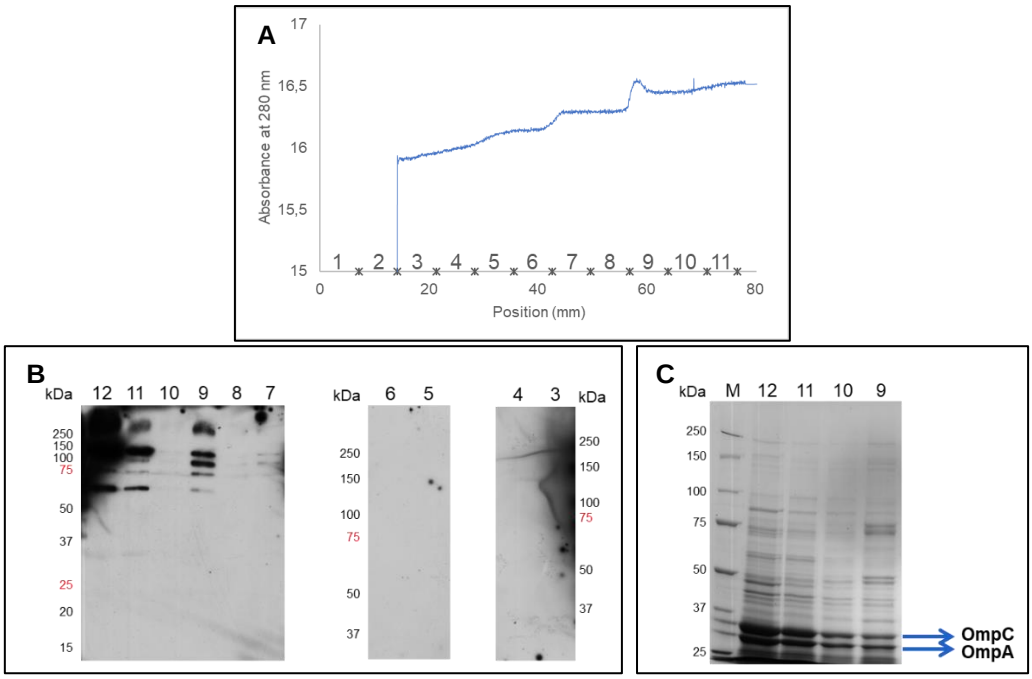

Figure 4.28: OptiPrep™ density ultracentrifugation eOMVs, derived from JC8031 after InaKtrBM3 production, under the control of the *araBAD* promotor. A: Absorbance at 280 nm of OptiPrep™ gradient after centrifugation for 3 h at 141 000 g. 1 ml fractions were collected from top to bottom and were numbered accordingly. B: Western blot analysis of the collected fractions after OptiPrep™ density ultracentrifugation. C: SDS-PAGE analysis of the InaKtrBM3-positive fractions. The intense gel bands in between 25 kDa and 37 kDa were cut out and analysed by MALDI-TOF MS. The upper band was identified as OmpC, the lower band as OmpA.

3.2.3 OMV display of self-sufficient CYP52A13

Using the self-sufficient CYP102A1 enzyme, OMVs were successfully obtained, containing the surface display construct InaKtrBM3. Although it was debated whether actual sOMVs were obtained, extraction of eOMVs was obtained in four different cases, i.e. in one of two deletion mutants JC8031 or *nlpI* 863 in combination with either the pMAL or pBAD vector. We showed that all four InaKtrCYPBMR constructs (with CYP being either CYPtr, CYPΔ32, CYPΔ25 or CYPnt) were surface displayed so we wondered whether these surface displayed enzymes would also be included in JC8031-derived vesicles, i.e. eOMVs and/or sOMVs. The four different constructs had a markedly different influence on the growth characteristics of JC8031. Expressing the CYPtr and CYPnt construct did not seem to greatly impair growth (OD_{600} = 6.26 and 5.00, respectively). However, expression of the CYPΔ32 construct led to a significant decrease of the final optical density (OD_{600} = 3.12). Upon expression of the CYPΔ25 construct, little growth was observed after IPTG induction (final OD_{600} = 0.687). Indeed, eOMV collection including either the CYPΔ32 or CYPΔ25 construct was not possible as the cells lysed upon EDTA incubation. The remaining collected OMV samples were analysed by western blot and SDS-PAGE (Figure 4.29A and B, respectively). In case of CYPtr and CYPnt, only the eOMVs, and not the sOMVs, contained the self-sufficient CYP and in these samples, the OM markers OmpC and OmpA, but not OmpF, were identified. The sOMVs of CYPΔ32 and CYPΔ25 both contained the self-sufficient CYP. As sOMVs only contained the surface displayed enzyme when growth is clearly impaired, these sOMVs were more likely to be the result of cell lysis and no true sOMVs were formed.

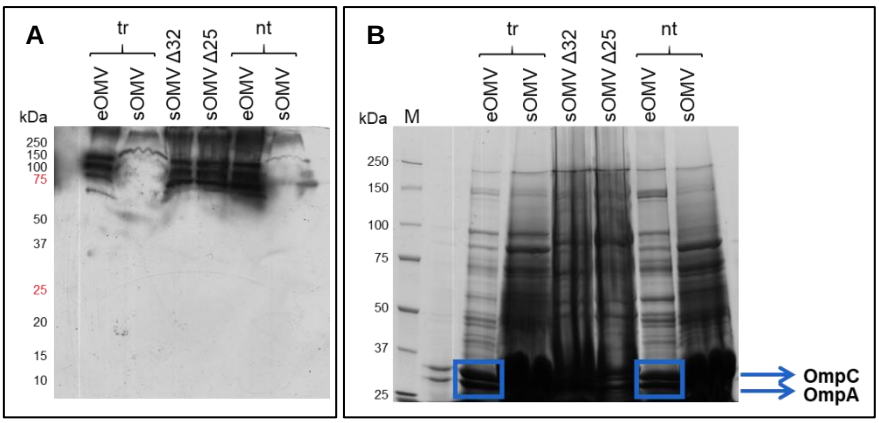

Figure 4.29: OMVs collected after InaKtrCYPBMR (CYP stands for either CYPtr, CYPΔ32, CYPΔ25 or CYPnt) production in JC8031. A: Western blot analysis. B: SDS-PAGE analysis. The gel bands indicated in the blue boxes were cut out and analysed by MALDI-TOF MS. Both upper bands were identified as OmpC, whereas both lower bands were identified as OmpA.

4 Discussion

4.1 InaKtr is a powerful tool for the surface display of CYPs

INPs have been reported in literature to be excellent anchors for the surface display of complex proteins and enzymes including cofactors have been successfully displayed [252], [253]. The value of INP as a surface display anchor was confirmed in this chapter, as all different constructs made, appeared at the surface of the *E. coli* cell. Especially for the construct where the N-terminal membrane anchor of CYP52A13 was not deleted, this was remarkable. Indeed it has been mentioned in literature that native class II CYPs can be incorporated into the membrane when provided with the signal sequence from PelB or OmpA. However, it was not specified whether the CYPs were incorporated in the IM or the OM and how the CYP was orientated (cytoplasm, periplasm or surface displayed) [242]. Here it thus seems that the native enzyme was integrated in the OM facing the extracellular space, based on the proteinase K accessibility assay. This raises questions: is the transmembrane anchor of CYP52A13 integrated in the OM or solvent exposed and how come the hydrophobic anchor was not retained in the IM? One might wonder if the proteinase K indeed was properly inhibited by the proteinase inhibitor cocktail and completely washed away before OM collection, and thereby not degrading intracellular and periplasmic proteins. All four constructs appear in either eOMVs or sOMVs, which strengthens the conclusion that the constructs are effectively integrated in the OM. However, again this conclusion should be taken with care, as it was not confirmed that the produced OMVs were derived from the OM alone and no IM was incorporated. The OMVs were not purified by OptiPrep™ density ultracentrifugation so it should also be taken into account that enzyme released upon cell lysis might adhere to the membranes, especially in case of the hydrophobic CYP52A13.

As powerful as the InaKtr anchor appears to be, optimization of the linker sequence fusing the anchor to the enzyme should still be addressed. The linker sequence was sufficient to retain enough enzyme at the surface of either the cell or the OMV in order to obtain an actively converting immobilized CYP, proven by InaKtrBM3. Unfortunately, a high amount of BM3 enzyme was found in the culture medium, indicating that the linker sequence is susceptible to proteolysis. Although the released BM3 was still active, this has implications for the costs of the catalytic process. Using either cells or OMVs gives the opportunity for easy catalyst recycling. However, when the enzyme is cleaved of the anchor, the activity of the recovered catalyst will significantly reduce at each recycling step.

4.2 The surface display of CYP52A13 chimers does not result in an oleic-acid converting biocatalyst

The molecular lego approach as described by Gilardi et al [273] was used in order to create a chimer between CYP52A13 and the reductase domain of CYP102A1. Three different truncations, as well as the nontruncated form of CYP52A13 were incorporated into such a chimer. Unfortunately, not one construct led to the conversion of oleic acid upon surface display. The same questions as put forward in previous chapter, can be asked here. Is BMR

capable of delivering electrons to the CYP domain? Is there uncoupling, leading to ROS formation, which might inhibit CYP? Can the substrate in fact bind in the active site? Is the protein yield too low for functional assays? These questions were already addressed in previous chapter and these issues will therefore not be discussed here.

It should be realized that the OMVs displaying the InaKtrCYPBMR constructs could not be properly assayed as these were only obtained at the very end of this PhD study. Park et al showed how the use of OMVs increased activity, which was ascribed to the higher enzyme:volume ratio [349]. If the enzyme shows an activity too low to be detected after surface display, activity might still be observed using OMVs due to this higher ratio. Of note, next to InaK, other INPs have been identified and it was reported that alternative INPs differ in their efficiency to target their recombinant protein passenger to the surface [361]. Also Autodisplay resulted in successful display of CYPs on the surface of *E. coli* [247], [249], [254]. If the problem lies within the low activity, alternative anchors might be tested as well in order to improve the yield of surface displayed CYPs thereby improving the activity of the whole-cell biocatalyst.

An additional factor in the surface display of CYP52A13 is the fact that the membrane environment of the OM is markedly different from the ER membrane. The OM of the *E. coli* envelope consists of an asymmetrical bilayered membrane, existing of a phospholipid inner leaflet and an outer leaflet containing lipopolysaccharides (LPS). It has been reported that the membrane environment can significantly influence the activity of CYP and its redox partner CPR [167], [168]. In this regard, it might be interesting to extract OMVs using detergents. Detergent extraction methods have been extensively used in the development of vaccines using OMVs. The extracted dOMVs contain considerably less LPS while bacterial OM proteins and other membrane components are integrated [348]. These dOMVs could therefore serve as a way to investigate the influence of LPS on the enzyme activity.

Also regarding the involvement of the membrane in the overall catalysis, it might be interesting to test several alternative surface display anchors. The surface anchor InaKtr used in this dissertation contained both the N-terminal and C-terminal unique domain of the INP and this N-terminal surface anchor is markedly longer than the natural CYP membrane anchor. In fact, using the N-terminal INP surface anchor alone proved to be sufficient for surface display [252]. Shorter INP surface anchors might therefore be tested to mimic the natural membrane anchor more closely. Alternatively, the INP central repeating domain and/or the linker between the surface display anchor and the CYP could be altered in order to enable membrane interaction, which can be beneficial for CYP catalysis. The Autodisplay surface display anchor is attached to the CYP C-terminus instead of the N-terminus. Investigating the surface display of self-sufficient CYPs, this approach could be tested as well.

4.3 A proof-of-concept was delivered, demonstrating the use of OMVs displaying a self-sufficient CYP as nanobiocatalysts

In this chapter, OMVs displaying an active self-sufficient CYP enzyme, i.e. CYP102A1 from *B. megaterium*, were successfully isolated. EDTA-extracted eOMVs were obtained in four different cases, i.e. using one of two deletion mutants JC8031 or *nlpI* 863 in combination with either the pMAL or pBAD vector. This confirms the study in which Van der Waterbeemd showed that OMV extraction with EDTA from a gram-negative bacterium, genetically modified resulting in a loosely attached OM, yields a higher amount of vesicles compared to the isolation of spontaneously formed sOMVs [363]. Spontaneously formed sOMVs were isolated only when the *tolRA* double deletion mutant JC8031 was used to express the InaKtrBM3-encoding gene under the control of the *tac* promotor in the pMAL-c4x vector. It became evident that obtaining sOMVs in a sufficient yield to observe activity, requires a very delicate equilibrium between OM instability and viability. The JC8031 strain shows to be particularly well suited for this purpose, but only in combination with the strong *tac* promotor. In the first experiments, the strain transformed with the pMAL construct showed good optical densities at the point of OMV collection. Unfortunately, every consecutive round of OMV production, starting from a glycerol stock made during the previous experiment, a decreased optical density was reached. At a certain point, eOMVs could no longer be extracted as the cells lysed upon contact with the EDTA-buffer. This exemplifies this delicate equilibrium. A construct under the control of the *araBAD* promotor did not result in a sOMV yield, sufficient to observe activity. This promotor is proposed to be used in order to better control expression, however this negatively affects the sOMV yield. Knocking out *nlpI* was reported to result in hypervesiculation without affecting the membrane integrity. Indeed hypervesiculation was observed, but only after overnight growth. However, this increased membrane integrity negatively affected the sOMV yield. It thus seems that the membrane instability really needs to be pushed to its limits and a strong promotor is required for sOMV production. This is further evidenced by the fact that only the growth impaired strains, surface displaying either the CYPΔ32 or the CYPΔ25 construct, produced sOMVs including the recombinant proteins, whereas in case of the two other constructs, eOMVs but not sOMVs contained the recombinant protein. One might therefore wonder if these are indeed all true sOMVs or if cell lysis occurs and results in released membrane fractions containing the surface displayed construct. Indeed, it must be noted that knock-out strains with disrupted cross-links, often showed a leaky membrane and Toyofuku et al reported that vesicles formed by cell lysis might be co-isolated in that case [352]. In order to further characterize the OMVs, OptiPrep™ density ultracentrifugation was performed, in the same way as described by Park et al [349], except that the gradient started from 45 % OptiPrep™ at the bottom instead of 35 % OptiPrep™. Unfortunately, the protocol did not result in complete purification. Although the gradient was formed with lower density solutions, it was still questioned whether Park et al indeed purified their OMVs completely as also they observed their OMVs to be mainly in the bottom fractions [349]. Kim et al [369], as well as Chutkan et al [371], report pure OMVs to be present in fractions which are lower in density than reported by Park et al. In order to purify and further characterize the formed OMVs, longer centrifugation times should be performed and/or the centrifugation speed should be increased.

Throughout this chapter, an enzymatic assay with the substrate 12-pNCA was used in order to measure the activity spectrophotometrically in case of InaKtrBM3 display. This proved to be a valuable alternative to GC-MS. However, it was observed that only a fraction of the added substrate seemed to be converted to p-nitrophenol. It must be noted that only hydroxylation of 12-pNCA at the ω-position will result in the chromophore p-nitrophenol. BM3 preferentially hydroxylates the subterminal positions ω-1, ω-2 and ω-3. Products resulting from hydroxylation at a position other than the terminal position will not be visible by spectrophotometry, explaining the low yield of p-nitrophenol. Using the F87A mutant, the conversion of 12-pNCA to p-nitrophenol was much higher so this this mutant might be used instead to enhance sensitivity [364]. Schwaneberg et al proposed that this assay might also be used for ω-hydroxylating CYP enzymes such as the CYP4 and CYP52 family [364]. The CYP52A13 chimeric constructs were tested for activity using this assay as well. However, the validity of this assay for ω-hydroxylating CYPs was questioned. Based on the relative bond strengths of subsequent C-H bonds in the hydrocarbon chain, the ω-1 position is preferentially hydroxylated over the terminal position and regiospecificity must be enforced through specific protein-substrate interactions. Inserting a phenoxy group at the hydrocarbon chain terminus was hypothesized to interfere with these specific interactions, based on the fact that CYP4 enzymes do not convert substrates where a phenyl group was located at the chain terminus [372]. In case of CYP52A21 [373] and CYP52A23 [53], it was hypothesized that a constricted access channel enables the regioselective hydroxylation at the ω-position. FAs were converted to halogenated derivatives with the halogen atom at the terminal position. It was observed that with increasing halogen atom size, access to the terminal carbon atom was impaired. Therefore, the aforementioned assay was not included in case of InaKtrCYPBMR display.

5 Conclusion

In this chapter, it could be shown that InaKtr is a powerful tool for the surface display of self-sufficient enzymes. Unfortunately, the surface display of CYP52A13, fused to BMR, did not result in an oleic acid converting biocatalyst. OMVs, displaying a self-sufficient enzyme, were successfully obtained, by either isolating sOMVs from the medium or extracting eOMVs from the cells. In case of CYP102A1 display, the vesicles showed to be active, thereby delivering a proof-of-concept for the use of OMVs as an *in vivo* CYP immobilization strategy for their use as nanobiocatalysts. The origin of the sOMVs, i.e. are the vesicles indeed spontaneously formed by blebbing of the OM or are these the result of cell lysis, requires further investigation.

Chapter 5
Recombinant production of a $NADP^+$-dependent formate dehydrogenase for cofactor regeneration

Delphine Devriese[1], Stijn De Waele[1], Bart Devreese[1]

[1]Laboratory of Microbiology – Protein Research Unit, Ghent University

Authors contribution: Delphine Devriese wrote the whole chapter and performed all experiments. Stijn De Waele designed the FDH construct for production in *P. pastoris*. Bart Devreese is the supervisor of the author.

Abstract

Many oxidoreductases show valuable activities for industrial applications. Unfortunately, their use is often limited due to their requirement for a cofactor. This is also the case for the enzyme of interest in this dissertation, i.e. CYP52A13 requires one mole NADPH for the ω-oxidation of one mole long-chain fatty acids to ω-hydroxy fatty acid and two extra moles for the overoxidation to the corresponding long-chain α,ω-dicarboxylic acid. As NADPH is an expensive cofactor, large-scale application of this system would not be feasible if the NADPH has to be added in stoichiometric amounts. For this reason, a NADPH-regeneration system was investigated. Formate dehydrogenase (FDH) was chosen as it is irreversible, requires an inexpensive co-substrate and the formed CO_2 is inert and can be easily removed. FDH from *Burkholderia stabilis* 15516 was selected as it has been reported to show NADP$^+$-specificity. In this chapter, a secretory approach was pursued in two different recombinant hosts, *P. pastoris* and *E. coli*, for simplified downstream processing. In *P. pastoris*, active FDH was purified from the medium. Unfortunately, stability issues limited its use. No active FDH could be obtained in the recombinant host *E. coli*.

1 Introduction

CYP monooxygenases are oxidoreductases, requiring a redox partner for the delivery of electrons from a hydride donor. This hydride is delivered by either NADH or NADPH. CYPs from class IV and class X are notable exceptions to this, as they do not use NAD(P)H [134]. The need for such cofactors is one of the major limitations in CYP-based biocatalysis [120], [122] and also affects the industrial implementation of other oxidoreductases. NAD(P)H is expensive and the addition of this cofactor in stoichiometric amounts is economically not feasible. Therefore, a cofactor regeneration system is required.

Formate dehydrogenases (FDHs) are a diverse set of enzymes, varying in quaternary structure, prosthetic groups and substrate specificity. One group exists of the NAD$^+$-dependent FDH enzymes, lacking metal ions or prosthetic groups and they are generally found as a dimer of two identical subunits. These enzymes catalyse the conversion of formate to CO_2, thereby reducing NAD$^+$ to NADH. The hydride ion is directly transferred from formate to the C4 atom of the nicotinamide without acid-base catalysis. The catalytic mechanism and the amino acid residues involved herein were extensively described in the past and are highly conserved in the FDH enzymes [374], [375].

FDH is of particular interest for cofactor regeneration on an industrial scale due to its many advantages, coinciding with prerequisites for low cost recycling. The FDH substrate formate is cheap and does not interfere with other enzymes. Only one enzyme has been reported so far to be inhibited by formate, that is xylitol reductase. Additionally, the reaction is irreversible. Furthermore, the formed end product CO_2 is easily removed and therefore, it does not interfere with purification processes [291]. A very successful application is the NADH recycling by FDH from *Candida boidinii* in the ton-scale production of *tert*-leucine [376]. Unfortunately, some disadvantages have limited the large scale industrial application of FDHs. Native enzymes have a low specific activity, a limited chemical and thermal stability and most enzymes specifically prefer NAD$^+$ instead of NADP$^+$ [377]. Many protein engineering studies have been done to improve these properties. Both rational design, for example the substitution of cysteine residues in order to improve chemical stability, and directed evolution approaches were reported. The plethora of FDH engineering studies has been reviewed in [378] and this was updated more recently in [379]. Although many improvements have been accomplished by protein engineering, this has not yet resulted in a NADP$^+$-dependent FDH for industrial application.

Glucose dehydrogenase (GDH) offers similar advantages, i.e. the glucose substrate is cheap and does not interfere with enzyme activities and the reaction is irreversible because the end product gluconolactone spontaneously hydrolyses to gluconic acid. GDH has the additional advantages of being highly active and more stable and accepts $NADP^+$. Therefore, this dehydrogenase is widely used, mostly in *in vivo* whole-cell bioconversion. Its use in *in vitro* application is limited due to the fact that the formed gluconic acid acidifies the reaction mixture and it needs to be purified from the main product. More recently, phosphite dehydrogenase (PTDH) has emerged as an interesting cofactor regeneration system, although protein engineering is required to increase catalytic activity, thermostability and $NADP^+$ specificity [291], [380], [381].

Hatrongjit and Packdibamrung identified the first native $NADP^+$-dependent FDHs in five species from the *Burkholderia cepacia* complex (BCC), where the FDH from *B. stabilis* showed the highest specific activity [382]. All NAD^+-dependent FDHs contain the conserved NAD^+ binding sequence G(A)XGXXG$(X)_{17}$D and the aspartic acid determines the NAD^+ specificity, by interacting with the 2'- and 3'-OH of the NAD^+ ribose ring [374]. In FDHs from the BCC species, the aspartic acid is replaced by a glutamine residue and it was shown that substitution of this residue with aspartic acid, changed the $NADP^+$ specificity to NAD^+ [382]. Using FDH from *B. stabilis*, further efforts were done to alter cofactor specificity of NAD^+-dependent FDH enzymes by protein engineering [377]. Alternatively, the sequence was used as a template for genome mining studies, resulting in the identification of two other $NADP^+$-dependent FDHs, i.e. the acidic, thermostable FDH from *Lactobacillus buchneri* NRRL B-30929 [383] and the acidic FDH from *Granulicella mallensis* MP5ACTX8 [384]. In our search for a NADPH regeneration system, we selected FDH from *B. stabilis* for recombinant production.

The recombinant production host *Pichia pastoris* was preferred, due to its many advantages for industrial applications as described before (Chapter 3 - 1). FDH production in a secreted form was pursued in order to simplify downstream processing. FDH from *C. boidinii* has been successfully produced in a secreted form in *P. pastoris* [385] and recently, the NAD^+-dependent *Chaetomium thermophilum* FDH was purified from the *P. pastoris* medium as well [386]. Next to *P. pastoris,* FDH production in the recombinant host *Escherichia coli* was investigated. Generally, FDH enzymes are produced in the cytoplasm, although surface display of FDH has been reported as well [387]. It was decided to direct FDH to the periplasm. This choice was again based on economic perspectives to reduce purification efforts. Indeed, leakage to the medium occurs in this system, which is enhanced by prolonged induction times. In the medium, less contaminating proteins are present, compared to the intracellular space, thus leading to simplified downstream processing [313].

2 Material and methods

2.1 Materials

All chemicals were purchased from Sigma-Aldrich, unless specified otherwise.

2.2 Strains and plasmids

For recombinant protein production in *P. pastoris*, the coding sequence for FDH from *Burkholderia stabilis* 15516 was synthesized, codon optimized and cloned in the vector pPICZαB in frame of the α-factor secretion signal in between restriction sites PstI and SalI by GenScript. The stopcodon was removed from the synthesized sequence in order to include the Histag at the C-terminus (see Addendum). FDH was produced in the *P. pastoris* strain NRRL-Y-11430 (kindly provided by Prof. Nico Callewaert (VIB, Ghent University)). The transformation was performed in the same way as described before (Chapter 3 - 2.5), except that PmeI was used for plasmid linearization. The plasmid was maintained in *E. coli* XL1-blue.

Similarly, the coding sequence was synthesized, codon optimized and cloned in the vector pET20b(+) in frame of the *pelB* signal sequence in between restriction sites XhoI and NcoI by GenScript for recombinant protein production in *E. coli*. Here as well, the stopcodon was removed in order to attach a C-terminal Histag (see Addendum). The FDH_pET20b(+) construct was transformed in *E. coli* competent DH5α (NEB) cells by electroporation for plasmid maintaining. The construct was transformed in *E. coli* competent BL21(DE3) (a kind gift from Prof. Savvas Savvides (VIB, Ghent University)) cells by electroporation for recombinant protein production. Transformants were selected in LB medium (1 % (w/v) NaCl (Merck), 0.5 % (w/v) yeast extract (Lab M) and 1 % (w/v) tryptone (Lab M)) including 100 µg/ml carbenicillin (Cb) (Gold Biotechnology). Alternative *E. coli* strains used for recombinant production were BL21-AI (Invitrogen) and the BL21(DE3) pLysS host (a kind gift from Prof. Savvas Savvides (VIB, Ghent University)).

2.3 Recombinant protein production in *Pichia pastoris*

Protein production was started with a 10 ml preculture in YPD medium with 100 µg/ml Zeocin™, incubated overnight at 28 °C and 250 rpm upon inoculation with a colony or with 1 ml from a glycerol stock. First a growth phase of 48 h was executed where *P. pastoris* was grown in BMGY medium (1 % (w/v) yeast extract (Lab M), 2 % (w/v) peptone (BD), 100 mM potassium phosphate buffer, pH 6, 1.34 % (w/v) YNB (Formedium) and 1 % glycerol (v/v)), inoculated with preculture (1/50), at 28 °C and 250 rpm. Subsequently, the cells were washed with BMMY medium (BMGY medium with glycerol replaced by methanol (VWR)), followed by an expression phase of 48 h in BMMY medium. The expression phase was carried out at 28 °C and 250 rpm whereby expression was induced with the addition of methanol to a final concentration of 1 % (v/v) every 12 h. Alternatively, cells were incubated at 16 °C and/or expression was induced with methanol to a final concentration of 0.5 % (v/v) methanol every 24 h or of 3 % methanol (v/v) every 12 h.

Small scale expression tests were carried out in 10 ml cultures in 50 ml birmboin tubes. Midscale expression tests refer to 500 ml cultures in 3 l baffled shake flasks.

2.4 Recombinant protein production in *Escherichia coli*

A preculture was set up, inoculating 5 ml LB including 100 µg/ml Cb in a 50 ml birmboin tube with a colony or 500 µl from a glycerol stock. The preculture was incubated overnight at 37 °C and 200 rpm. The next day, cells were grown in LB including 100 µg/ml Cb, inoculated with 1/100 volume preculture. Cells were grown at 28 °C and 200 rpm until an OD_{600} of 0.8 was reached. At this point, expression was induced by adding isopropyl-β-D-thiogalactoside (IPTG) (Gold Biotechnology) to a final concentration of 0.4 mM. An overnight expression phase was carried out in order to promote protein leakage to the medium. Exploratory experiments were carried out in 50 ml cultures using 500 ml shake flasks and alternative expression conditions were tested, as denoted in the results section. For protein purification purposes, 300 ml cultures were set up in a 3 l shake flask.

The same method was used in case of the pLysS host and the BL21-AI strain, with some modifications. In case of the pLysS host, 25 µg/ml chloramphenicol was included together with Cb. For the BL21-AI expression host, induction was performed at an OD_{600} of 0.4 with arabinose to a final concentration of 0.2 %. Autoinduction was executed as described by Studier et al [360], [370]. After the expression phase, samples were taken and mixed with 2X Laemmli for western blot analysis. These samples are referred to as total cell protein or TCP.

2.5 *Escherichia coli* cell fractionation

2.5.1 Target protein verification

Upon the first exploratory expression tests, the cultures were collected and fractionated in order to verify where de produced FDH resided. To this end, cells were pelleted by centrifugation at 10 000 g and 4 °C for 10 min. A 1 ml medium sample was taken and concentrated by DOC-TCA precipitation (as described in Chapter 3 - 2.8). The precipitated protein was resuspended in Laemmli for western blot analysis. The periplasmic fraction was collected by means of osmotic shock. Therefore, cells were first resuspended in 30 mM Tris-HCl, pH 8, containing 20 % sucrose and 1 mM EDTA. The resuspended cells were gently stirred for 10 min at room temperature. The cells were subsequently pelleted at 10 000 g and 4 °C for 10 min. The pellet was subsequently resuspended in ice-cold 5 mM $MgSO_4$ and gently stirred for 10 min on ice. Next, the shocked cells were pelleted and a 1 ml supernatant sample was taken for DOC-TCA precipitation. The precipitate was then resuspended in Laemmli buffer. In a following step, cells were resuspended in ice-cold 20 mM Tris-HCl, pH 7.5, including protease inhibitor and lysed by sonication. The soluble lysate was separated from the insoluble lysate by centrifugation. A sample of the soluble lysate was taken and mixed with 2X Laemmli buffer for western blot analysis. The insoluble pellet was washed with 20 mM Tris-HCl, pH 7.5 and then resuspended in a 1 % (w/v) SDS solution. Solubilization was enhanced by heating and vigorous mixing. Subsequently, a sample was mixed with 2X Laemmli buffer.

2.5.2 Sample preparation protein purification

After a first midscale expression test, cells were pelleted by centrifugation at 10 000 g and 4 °C for 10 min and the medium was collected for IMAC purification. The cell pellet was resuspended in 30 mM Tris-HCl, pH 8, including 1 mM EDTA and 20 % (w/v) sucrose. The suspension was divided in two in order to test two different methods for periplasmic protein collection. The first half was subjected to osmotic shock. To this end, the suspension was first incubated on ice for 1 h with gentle agitation. Cells were pelleted by centrifugation and the supernatant was saved. The collected cells were subsequently resuspended in a hypotonic solution containing 5 mM $MgSO_4$. A second incubation step on ice for 1 h with gentle agitation was performed. The cells were spinned down and the two supernatant fractions were pooled and diluted 1/5 with IMAC binding buffer for subsequent protein purification. To the second half of the resuspended cells, 1 mg/ml lysozyme was added and incubated on ice for 1 h with gentle agitation. Cells were pelleted and the supernatant was diluted 1/10 with IMAC binding buffer, followed by protein purification.

A second midscale expression test was performed, after which the culture was divided in three equal parts. To the first part, Tris, NaCl and $MgSO_4$ was added to a final concentration of 100 mM, 300 mM and 2 mM, respectively. This suspension was incubated for 4 h on ice with gentle agitation in order to increase protein leakage to the medium. Centrifugation was performed at 10 000 g and 4 °C for 10 min and the supernatant was saved. The pelleted cells were resuspended in 100 mM Tris-HCl, pH 7.5, 500 mM NaCl and 2 mM $MgSO_4$. Incubation was carried out for 2 h on ice with gentle agitation. After centrifugation, the two supernatant fractions were pooled and FDH was purified by IMAC and SEC. The second part was subjected to osmotic shock as described above. The two supernatant fractions were pooled and a buffer exchange, using the Amicon® Ultra-15 centrifugal filter devices (Merck) with a 10 kDa cutoff, was performed to the IMAC binding buffer in order to remove EDTA. This sample, together with the collected medium (supplemented with NaCl to a final concentration of 300 mM) of this second part was further purified by IMAC and SEC. The third part of the culture was lysed. Cells and medium were separated by centrifugation at 10 000 g and 4 °C for 10 min. The medium was saved and NaCl was added to a final concentration of 300 mM. Cells were resuspended in 20 mM Tris-HCl, pH 7.5, including 500 mM NaCl and protease inhibitor. Cells were lysed by sonication. After pelleting the cell debris, the soluble lysate was pooled with the medium and purified by IMAC and SEC.

2.6 SDS-PAGE and western blot

The same protocol as described before (Chapter 3 - 2.9) was applied. As FDH contains a C-terminal Histag, only the anti-His(C-term)-HRP antibody (R931-25, Invitrogen) was used.

2.7 Protein purification

2.7.1 IMAC purification

Immobilized Metal Affinity Chromatography (IMAC) purification was performed in the same way as described in Chapter 3 - 2.10. Alternatively, 5 % glycerol was included in the all buffers.

2.7.2 SEC purification

Size exclusion chromatography (SEC) purification was performed with the column HiLoad™ 16/600 Superdex™ 200 prep grade (Cytiva), mounted on the ÄKTA purifier (Cytiva). The IMAC fractions containing FDH were injected on the SEC column after equilibration of the column with the a buffer consisting of 50 mM Tris-HCl, pH 7.5, and 150 mM NaCl. Upon injection, the same buffer was used for elution. Fractions of 1 ml were collected and those belonging to the same peak were pooled and concentrated using the Amicon® Ultra-15 centrifugal filter devices (Merck). Alternatively, 5 % glycerol was included in this SEC buffer.

2.8 Trypsin digestion and MALDI-TOF MS

The same protocol as described before (Chapter 3 - 2.12) was applied. Identification was done with Mascot using either the *P. pastoris* NRRL-Y-11430 protein database or the *E. coli* (strain K12) protein database, downloaded from Uniprot.

2.9 Determination of protein concentration

Protein concentrations were determined with the Thermo Scientific™ Coomassie (Bradford) Protein Assay Kit. The measurement was done using the Bio-Rad Microplate Reader model 680.

2.10 Enzymatic assay

FDH activity was measured spectrophotometrically by monitoring the formed NADPH at 340 nm (ε = 6220 M^{-1} cm^{-1}) using the Bio-Rad Microplate Reader model 680. The reaction was carried out in a volume of 250 µl consisting of 200 mM potassium phosphate buffer, pH 7.4, 1 mM NADP$^+$ and 0.1 mM sodium formate. The reaction was initiated by simultaneous addition of the substrate sodium formate and was incubated for 30 min at 30 °C.

3 Results

3.1 Recombinant FDH production using *Pichia pastoris* as host

The complete coding sequence of *fdh* from *B. stabilis* 15516 was codon optimized and cloned in the vector pPICZαB in frame of the α-factor secretion signal, inserting a C-terminal Histag for IMAC purification. As mentioned before, expression in *P. pastoris* and secretion to the medium was preferred as this already serves as a first purification step because *P. pastoris* secretes little endogenous proteins. Following transformation, three *P. pastoris* colonies were picked for a first small scale expression test. A sample of the medium was taken, concentrated and analysed by western blot (Figure 5.1). In parallel, cells were lysed and a sample of the lysate was taken along. In the medium, a clear signal was observed for all three colonies at the expected MW of 43 kDa. No signal was seen in the lysate, indicating that the protein production and secretion is efficient and no intracellular degradation occurs.

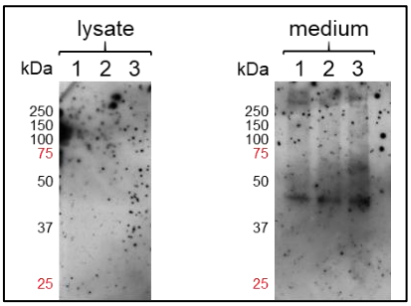

Figure 5.1: Western blot of lysate and concentrated medium samples after small-scale expression test of three colonies picked after *fdh* transformation in *P. pastoris*. FDH (43 kDa) was successfully secreted to the medium. This secretion appeared to be efficient as no signal was observed in the lysate.

After the small-scale expression test, midscale expression tests were performed, followed by FDH purification from the medium in two steps. Firstly, FDH was purified by IMAC, and secondly, SEC was performed as the polishing step. In a first batch of cultures, methanol concentrations were varied during the expression phase. Standard practice in the lab is the addition of methanol to a final concentration of 1 % for inducing recombinant protein production and additional methanol supplementation every 12 h. Alternatively, the methanol concentration was lowered to a final concentration of 0.5 % for the induction of recombinant protein production and additional supplementation was only done once after 24 h. As methanol metabolism requires a high oxygen supply, oxygen depletion might result in a reduced yield when adding 1 % methanol every 12 h during the expression phase. Lowering the methanol concentration might thus reduce oxygen limitations. Lastly, the methanol concentration was increased to a final concentration of 3 % for the induction of recombinant protein production, and supplemented every 12 h during the expression phase. Higher methanol concentrations have been reported to result in less byproduct formation and to increase recombinant protein production selectivity [388]. The OD_{600} at the point of collection after 0.5 % methanol addition was 57.3, after 1 % methanol addition 56.6 and after 3 % methanol addition 55.6. Based on these optical densities, the different methanol supplementation strategies did not greatly influence the final cell density. Medium was collected and subsequently purified by IMAC (Figure 5.2). It must be noted that during sample loading after both 1 % and 3 % methanol

addition, precipitation issues occurred. Looking at the three different chromatograms, using a limited methanol supplementation strategy seemed to result in a slightly higher yield as the 100 mM imidazole elution peak has the highest absorbance. However, together with FDH, contaminating proteins co-eluted with 100 mM imidazole, as seen in the SDS-PAGE analysis of the different elution peaks (Figure 5.3) and these seemed to be relatively more present in case of 0.5 % methanol supplementation. A small amount of FDH eluted at a higher imidazole concentration of 250 mM. The SDS-PAGE analysis of the IMAC elution peaks collected from the medium obtained after 3 % methanol addition did not show any signal and is therefore not shown. In this case, all protein thus appeared to degrade during concentration and buffer exchange. Following IMAC, both 100 mM and 250 mM fractions were combined and further purified by SEC (Figure 5.4). Indeed, no major peaks were seen in case of 3 % methanol addition, confirming that all protein was lost after IMAC (data not shown). In the other two cases (Figure 5.4A and B), several peaks appeared in the chromatogram and all were collected for SDS-PAGE analysis (data not shown). However, no protein was visualized and western blot analysis, being more sensitive, was needed to see which elution peaks were in fact FDH (Figure 5.4C). Purity was therefore not assessed.

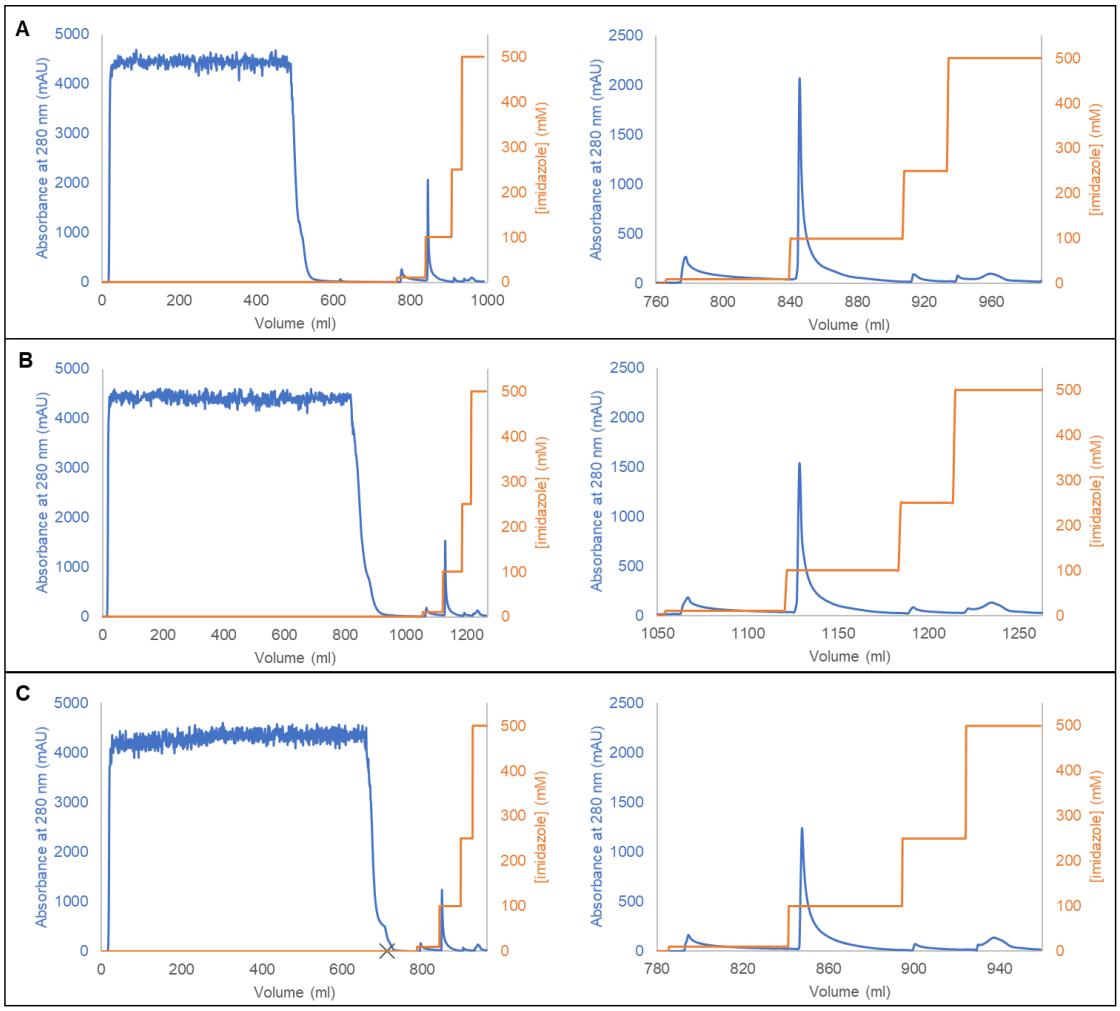

Figure 5.2: IMAC purification of FDH after midscale expression tests in *P. pastoris*. Three different methanol supplementation strategies were tested at 28 °C. A: 0.5 % MeOH every 24 h. B: 1 % MeOH every 12 h. C: 3 % MeOH every 12 h. For all purifications the complete chromatogram is shown on the left. The right chromatogram shows the elution profile in more detail.

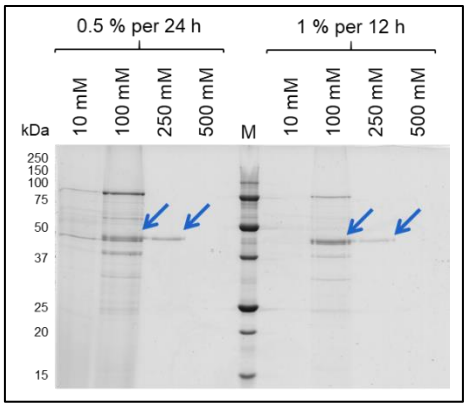

Figure 5.3: SDS-PAGE analysis of IMAC elution peaks, desalted and concentrated after FDH (43 kDa) purification (Figure 5.2). Three different methanol supplementation strategies were tested during the expression phase, i.e. 0.5 % MeOH every 24 h, 1 % MeOH every 12 h and 3 % MeOH every 12 h. The last one is not shown as no bands were seen on the Coomassie-stained gel. The arrow indicates the cut out gel bands, which were identified to be FDH by MALDI-TOF MS.

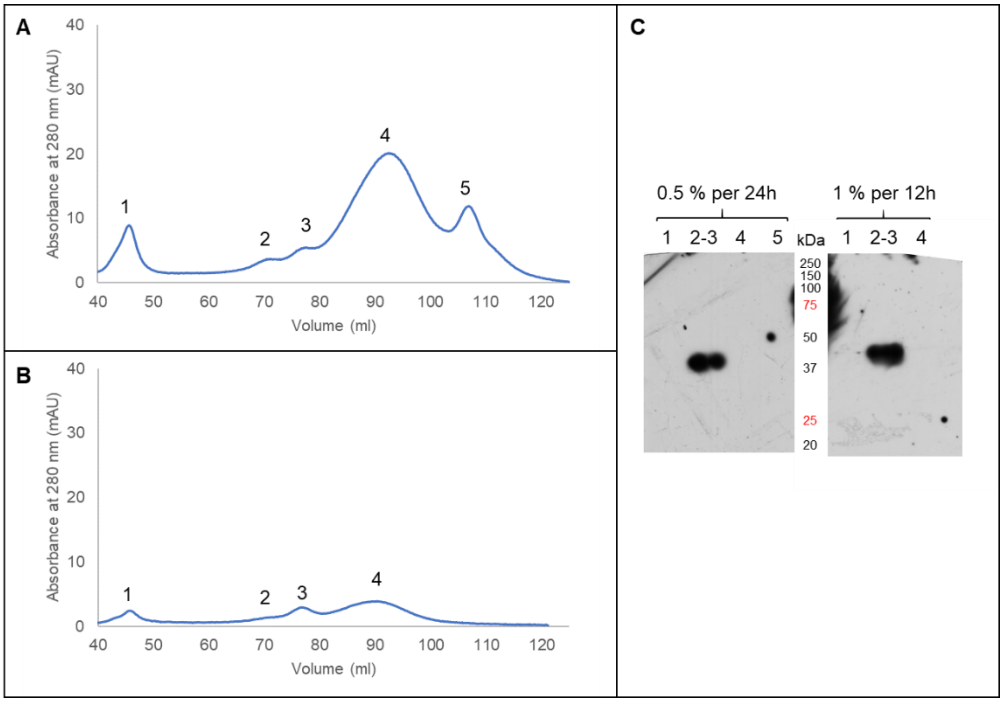

Figure 5.4: SEC purification of FDH (43 kDa) after midscale expression tests in *P. pastoris* and a first IMAC purification step (Figure 5.2 and Figure 5.3). Three different methanol supplementation strategies were tested at 28 °C, and only in two cases, FDH was retrieved after IMAC. A: 0.5 % MeOH every 24 h. B: 1 % MeOH every 12 h. C: western blot analysis.

In a next batch of cultures, the temperature during the expression phase was reduced to 16 °C instead of 28 °C, a common strategy to improve the yield of properly folded recombinant protein. By reducing the temperature during the expression phase, the metabolic activity is slowed down, leaving more time for the recombinant protein to fold correctly [339]. Both the 0.5 % and 1% methanol supplementation strategy as described above were tested in parallel. At the point of collection, an OD_{600} of 54 was measured in case of 0.5 % methanol, and of 50.7 in case of 1 % methanol addition. A minor decrease compared to expression at 28 °C was thus observed. The medium was collected and first purified by IMAC (Figure 5.5). Of note, precipitation issues again occurred in case of 1 % methanol addition. Therefore, the sample loading had to be interrupted, thereby losing the data. For this reason, the chromatogram in

Figure 5.5B (left) is incomplete. Based on the absorbance of the elution peak at 100 mM imidazole, the yield after expression at 16 °C using 1 % methanol supplementation per 12 h resulted in the highest yield, although many contaminants were observed in this fraction as well. In both cases, the 250 mM imidazole elution step also contained FDH and the two elution peaks were combined for further SEC purification (Figure 5.7). A smaller peak was seen around an elution volume of 76 ml. Following this smaller elution peak, a broader peak followed with a higher absorption maximum. Based on the subsequent SDS-PAGE analysis (Figure 5.7C), most FDH eluted around 76 ml. The broader peak contained FDH as well, together with lower MW bands and also a smear was observed. Of note, the protein samples were more concentrated than after the 28 °C expression experiments so the gel band intensities cannot be directly compared. It became evident during these experiments that the recombinantly produced and secreted FDH was unstable. Many precipitation issues occurred during IMAC sample loading and smearing was observed on the Coomassie-stained gels after SDS-PAGE.

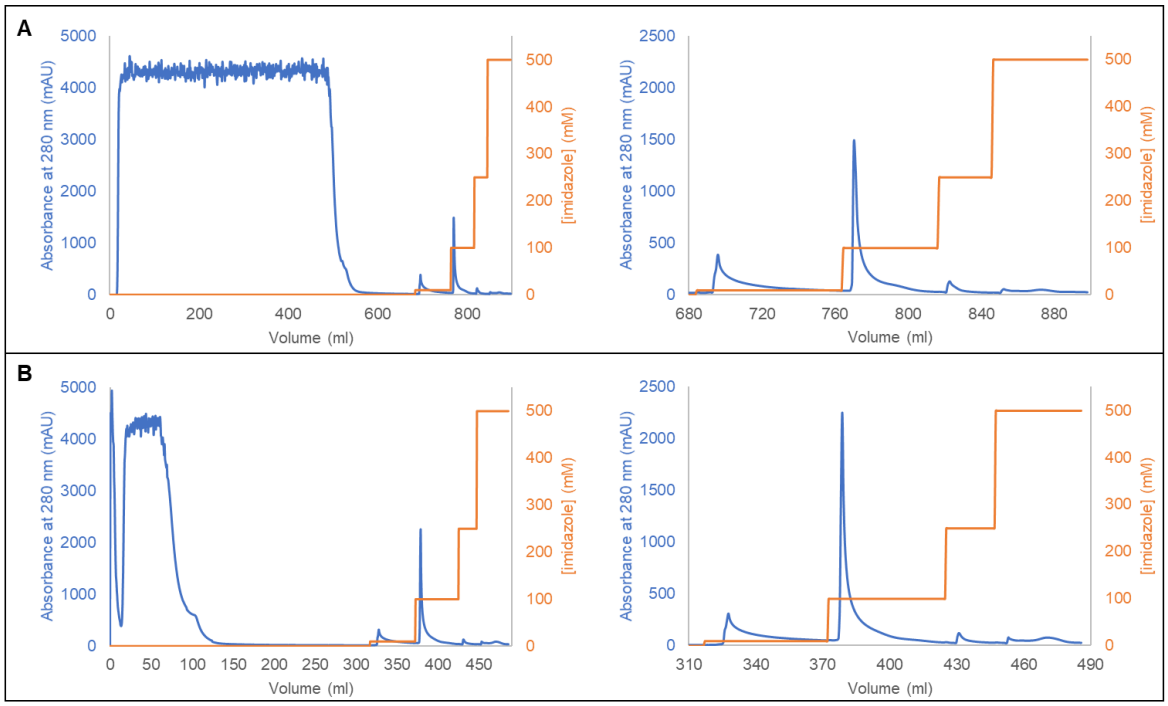

Figure 5.5: IMAC purification of FDH after midscale expression tests in *P. pastoris*. Two different methanol supplementation strategies were tested at a lower expression temperature of 16 °C. A: 0.5 % MeOH every 24 h. The complete chromatogram is shown on the left. The right chromatogram shows the elution profile in more detail. B: 1 % MeOH every 12 h. The complete chromatogram is shown on the left. Problems occurred during sample loading, leading to loss of data. Therefore the left chromatogram shows misses part of the sample loading data. The right chromatogram shows the elution profile in more detail.

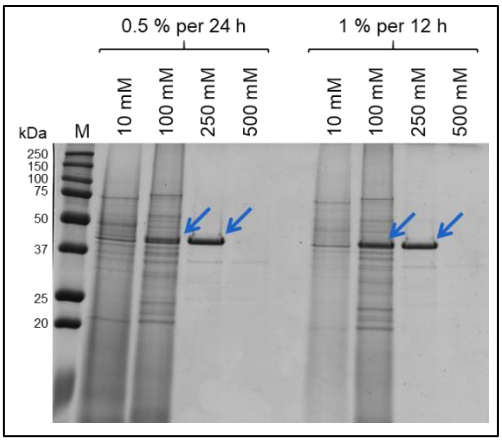

Figure 5.6: SDS-PAGE analysis of IMAC elution peaks, desalted and concentrated after FDH (43 kDa) purification (Figure 5.5). Two different methanol supplementation strategies were tested at a lower expression temperature of 16 °C, i.e. 0.5 % MeOH every 24 h and 1 % MeOH. The arrow indicates the cut out gel bands, which were identified to be FDH by MALDI-TOF MS.

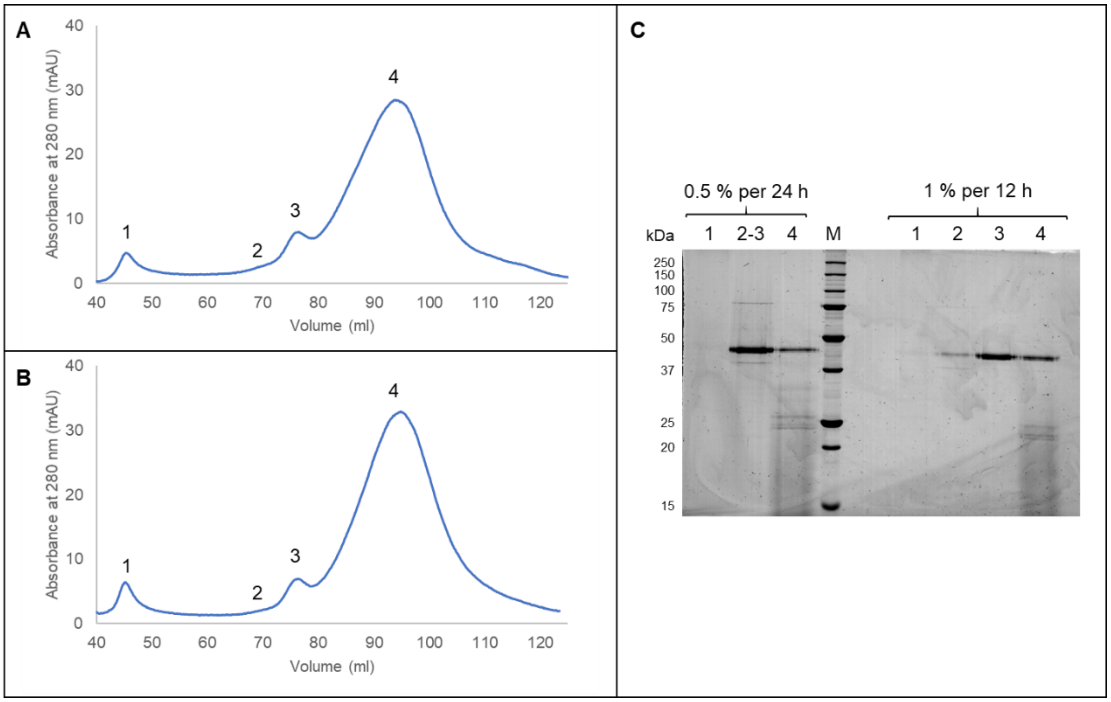

Figure 5.7: SEC purification of FDH (43 kDa) after midscale expression tests in *P. pastoris* and a first IMAC purification step (Figure 5.5 and Figure 5.6). Two different methanol supplementation strategies were tested at a lower expression temperature of 16 °C. A: 0.5 % MeOH every 24 h. B: 1 % MeOH. C: SDS-PAGE analysis.

In order to address the stability issues, encountered during midscale expression experiments, the stabilizing effect of glycerol was investigated. 5 % glycerol was added to the buffers, used for IMAC, SEC and storage. In a first experiment, the activity was measured after a first IMAC purification. Samples which were purified excluding glycerol and including glycerol in the buffers, were assayed in parallel (Figure 5.8A and B, respectively) after a first freeze-thaw cycle, and retested after a second freeze-thaw cycle. The addition of 5 % glycerol led to an increased stability as the activity was much more retained upon a second freeze-thaw cycle.

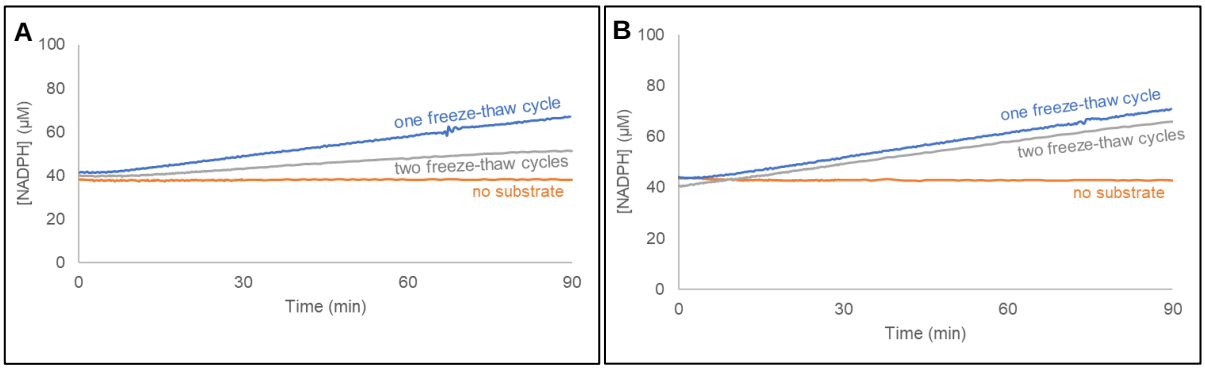

Figure 5.8: Activity assay IMAC purified FDH. NADPH formation was measured at 340 nm and the NADPH concentration was calculated with an extinction coefficient $\varepsilon_{340\,nm} = 6.22 \times 10^3\ M^{-1}\ cm^{-1}$. A: No glycerol was added to the purification and storage buffers. B: Glycerol was added to the buffers. The activity was retained when glycerol was included.

As the activity was retained upon glycerol addition, it was wondered if glycerol addition would lead to less protein loss during purification due to instability and thereby increasing the yield. In the IMAC chromatogram (Figure 5.9), no big differences in absorbance maxima were seen and subsequent SEC (Figure 5.11A) did not show significant absorbance increase either. The SDS-PAGE analysis after both purification steps (Figure 5.10 and Figure 5.11B, respectively) show that only little FDH was obtained after IMAC and that is was lost after SEC.

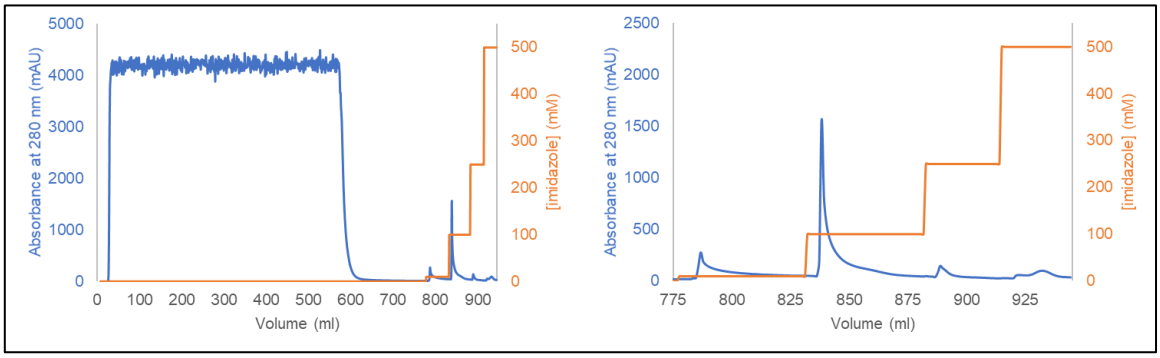

Figure 5.9: IMAC purification of FDH after midscale expression tests in *P. pastoris*, including glycerol in the used buffers. The complete IMAC chromatogram is shown on the left. The right chromatogram shows the elution profile in more detail.

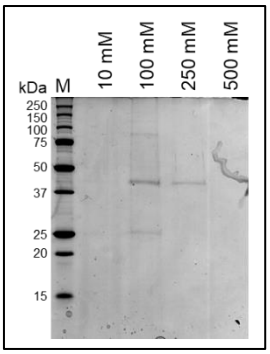

Figure 5.10: IMAC purification of FDH (43 kDa) after midscale expression tests in *P. pastoris*, including glycerol in the used buffers. SDS-PAGE analysis of IMAC elution peaks, desalted and concentrated after FDH purification (Figure 5.9)

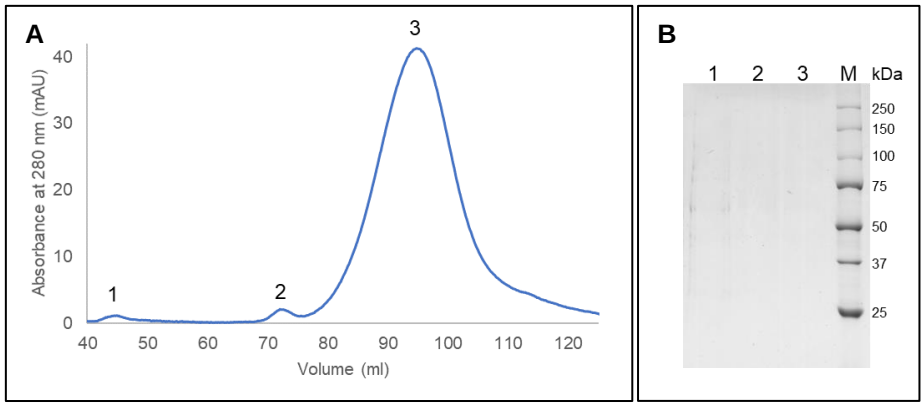

Figure 5.11: SEC purification of FDH (43 kDa) after midscale expression tests in *P. pastoris* and a first IMAC purification step (Figure 5.9 and Figure 5.10). A: SEC chromatogram. The numbers indicate the collected peaks. B: SDS-PAGE analysis of the SEC elution peak after concentration. No protein was recovered.

3.2 Recombinant FDH production using *Escherichia coli* as host

Recombinant FDH production and secretion in *P. pastoris* did not lead to great successes. A switch was therefore made to the alternative production host *E. coli*. It was chosen to clone *fdh* in frame of the *pelB* signal sequence of the pET20b(+) vector, thereby sending FDH to the periplasm. Prolonged induction times were applied in order to enhance leakage to the medium, leading to a simplified FDH purification process. A first exploratory expression test was performed at 37 °C, inducing recombinant expression with 0.4 mM IPTG, and a prolonged induction phase of 16 h was executed. A non-induced control was taken along. After expression, the cells were collected and the different cell fractions, i.e. periplasm, soluble lysate and the insoluble lysate (containing cell debris, unlysed cells and inclusion bodies (IB)) were isolated for western blot analysis, together with a concentrated medium sample (Figure 5.12). FDH was expressed in both the induced culture, as well as in the non-induced control but was only found in the insoluble lysate. As no soluble protein was present, it was concluded that the produced FDH was mainly present in inclusion bodies (IBs).

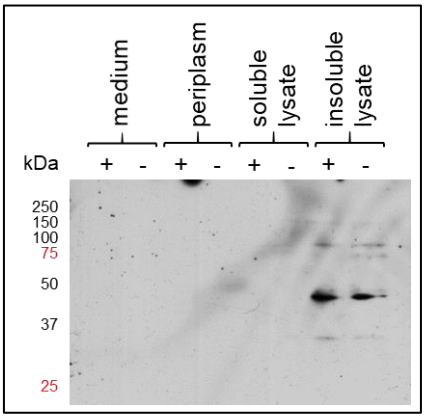

Figure 5.12: Western blot analysis of the collected cell fractions, that is the concentrated medium, periplasm, soluble lysate and insoluble lysate, after FDH (43 kDa) production in BL21(DE3) at 37 °C. +:induced culture, -:non-induced control.

In order to circumvent IB formation, the temperature was lowered in order to slow down protein production and allowing more time to properly fold. On the one hand, the temperature was lowered to 16 °C during the expression phase, after a growth phase at 37 °C where glucose was added to reduce leaky expression. On the other hand, the culture was incubated at 28 °C during both growth and expression. In parallel, a lower IPTG concentration for the induction of recombinant protein production was tested. The different cell fractions were collected and the western blot is shown in Figure 5.13. Incubation of the culture at 28 °C during both growth and expression resulted in the desired leakage to the medium, with induction using 0.4 mM IPTG resulting in a seemingly higher yield. However, leaky expression still occurred and in the soluble lysate, a signal at a higher MW is seen as well. Incubation at 37 °C including glucose during growth, followed by an expression phase at 16 °C showed to result in production of soluble FDH. In this case, the non-induced control seemed to result in the highest yield. Leakage to the medium was not obtained, although there seems to be a signal at a higher MW in the medium of the non-induced control. Of note, the optical densities obtained after growth at 37 °C, followed by expression at 16 °C, were lower than after incubation at 28 °C. In the first

case, the OD_{600} of the non-induced control was 2.685, induction with 0.1 mM IPTG resulted in an OD_{600} of 2.475 and induction with 0.4 mM resulted in an OD_{600} of 1.595. In the second case, the OD_{600} values were more similar, being 3.72, 2.9 and 3.34, respectively. Especially inducing with 0.4 mM IPTG thus resulted in a reduced cell density in the first case.

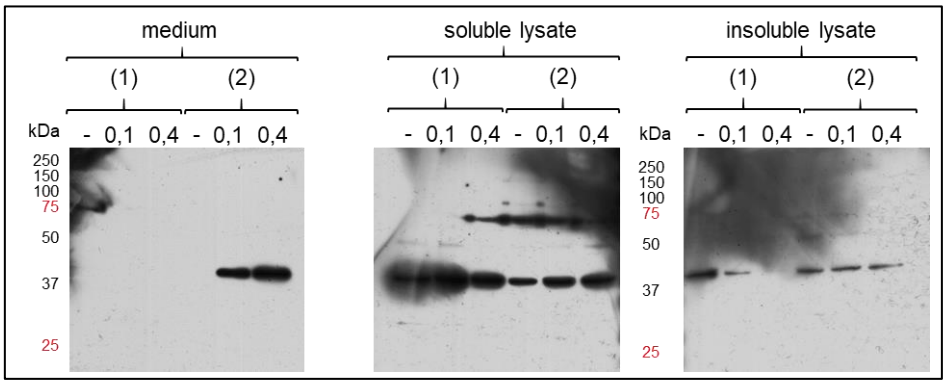

Figure 5.13: Western blot analysis of the collected cell fractions, that is the concentrated medium, soluble lysate and insoluble lysate, after FDH (43 kDa) production in BL21(DE3) at (1) 16 °C, where first the cells were grown at 37 °C, including 1 % glucose in the medium or (2) at 28 °C, where the cells were also grown at 28 °C. Induction of recombinant expression was performed with either 0.1 mM or 0.4 mM IPTG and a non-induced control was taken along in parallel (-).

After incubation at 28 °C during both the growth and expression phase, FDH was successfully recovered in the medium. Therefore, a larger scale culture was set up, followed by IMAC purification (Figure 5.14). Cells were pelleted and the medium was collected for IMAC purification. After resuspension, the cells were split in two in order to test two different protocols for FDH extraction from the periplasm. On the one hand, FDH was extracted from the periplasm by means of osmotic shock. On the other hand, lysozyme was used. Eluting with 100 mM imidazole, the highest absorption peak was observed, as seen during FDH purification from the medium after recombinant production in *P. pastoris*. However, the 100 mM imidazole elution peak still contained many contaminating proteins in case of FDH purification from the medium, as seen by SDS-PAGE, and no clear visible band around 43 kDa appeared. Both periplasmic extraction protocols resulted in distinct bands around 43 kDa. These bands were cut out for MALDI-TOF MS analysis. Unfortunately, this could not confirm the presence of FDH but instead, the bands appeared to contain the elongation factor Tu, a contaminant often encountered, which has a similar MW as FDH. A western blot analysis was therefore needed to find out if FDH was actually produced and purified. In the elution peak samples after FDH purification from the medium, no bands at the expected MW of 43 kDa were found. In both the 10 mM and 100 mM imidazole elution peaks, a band at a higher MW was observed instead. In the 100 mM imidazole elution peak, a band around 25 kDa was present as well, indicating degradation. In the elution peak samples after FDH purification from the periplasmic fraction, retrieved after osmotic shock, a signal at the right MW of 43 kDa appeared, however, eluting with the 10 mM imidazole washing step. No FDH was found after lysozyme treatment so osmotic shock was the periplasm extraction method of choice.

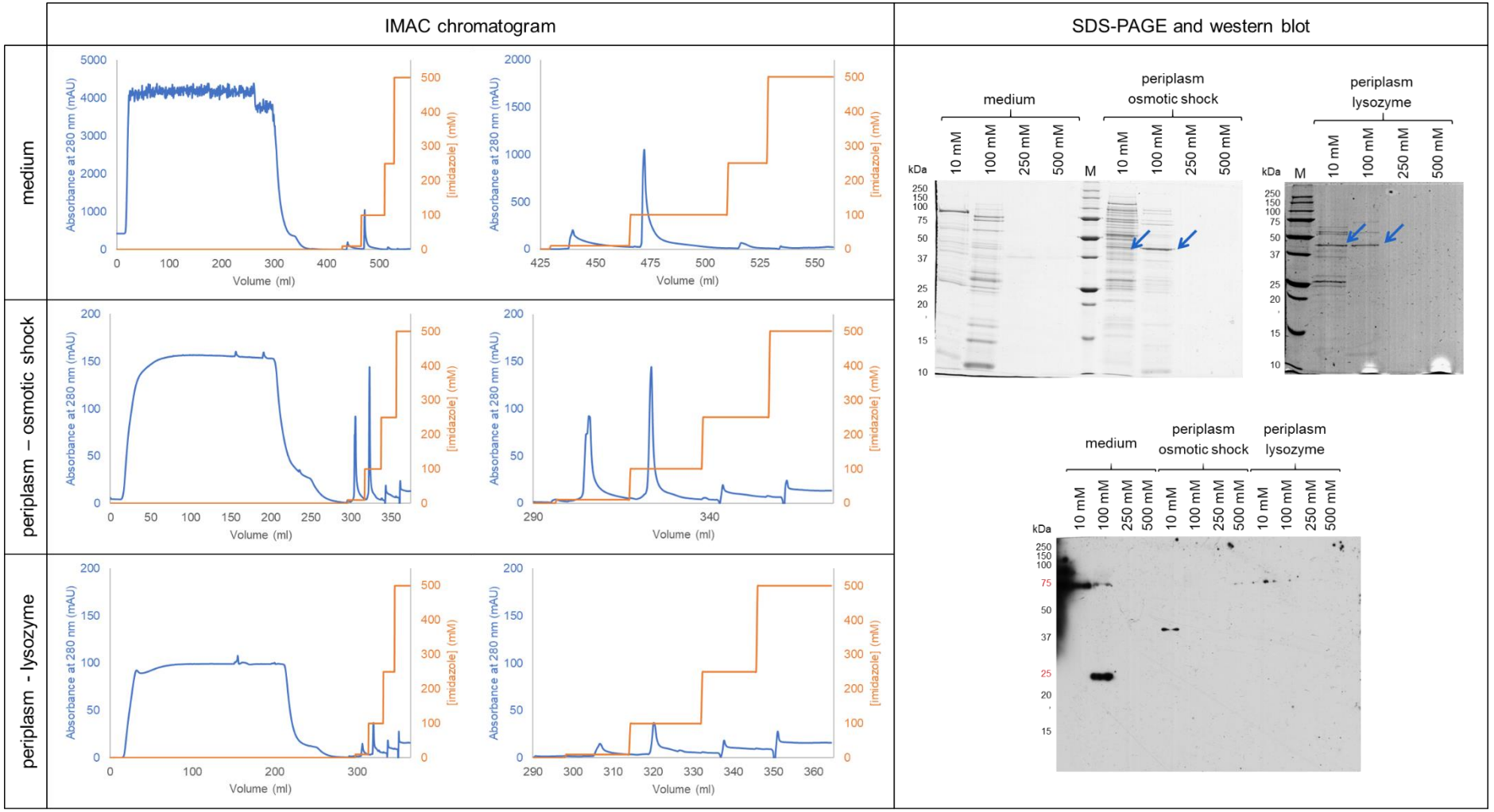

Figure 5.14: IMAC purification of FDH (43 kDa) after recombinant production in *E. coli*. The medium was collected and purified, as well as the periplasm. Cells were split in two and two different periplasm extraction protocols were tested, i.e. osmotic shock and lysozyme treatment. For all purifications, the complete chromatogram is shown on the left. The right chromatogram shows the elution profile in more detail. All elution peaks were analysed by SDS-PAGE and western blot, shown in the right column. The blue arrow indicates the cut out gel bands, which were identified to be elongation factor Tu by MALDI-TOF MS.

For FDH extraction from the periplasm, EDTA was included, which might interfere with IMAC purification. In both cases, the final periplasm extract was diluted in order to decrease the EDTA concentration to 0.1 mM, which is ten times lower than the limit mentioned in the HisTrap™ HP manual (1 mM). Moreover, Mg^{2+} was included to counteract EDTA. Despite the low concentration and inclusion of Mg^{2+}, it was hypothesized that the EDTA presence could still interfere with proper FDH purification using IMAC. Therefore, a buffer exchange to the IMAC binding buffer was performed after osmotic shock. In addition, two other approaches were tested in order to check which method resulted in the highest yield versus the highest purity. Firstly, the cells were incubated with Tris in order to increase the permeability of the OM, promoting FDH leakage to the medium. A second alternative was to lyse the cells completely in order to check whether the increased yield (as leakage and osmotic shock might not release all protein produced) can make up for the decreased purity. One 300 ml culture was divided in three 100 ml cultures after expression to make sure the yield difference indeed was the result of the FDH collection method. The medium was included as well in the periplasm extract and the lysate, as a prolonged induction phase was executed. The IMAC purification results are shown in Figure 5.15. In the Coomassie-stained gel of all three purifications, a clear band is visible around 43 kDa, especially when purifying the lysate. Again, gel bands were cut out and analysed by MALDI-TOF MS and here as well, FDH was not found, and these were mainly identified as elongation factor Tu. A subsequent western blot was executed, showing that very little FDH was purified. Only in the 10 mM elution peak of the FDH purification from the lysate, there appears a very light signal around the right MW of 43 kDa. This is also the case after medium leakage, although a more intense band is present at a higher MW. Now, a sample was taken of the flow through in order to see whether FDH was lost due to binding issues with the IMAC column. The sample was concentrated and analysed by western blot (Figure 5.16). No FDH was found so no FDH was lost via the flow through and it was concluded that the FDH yield was very low.

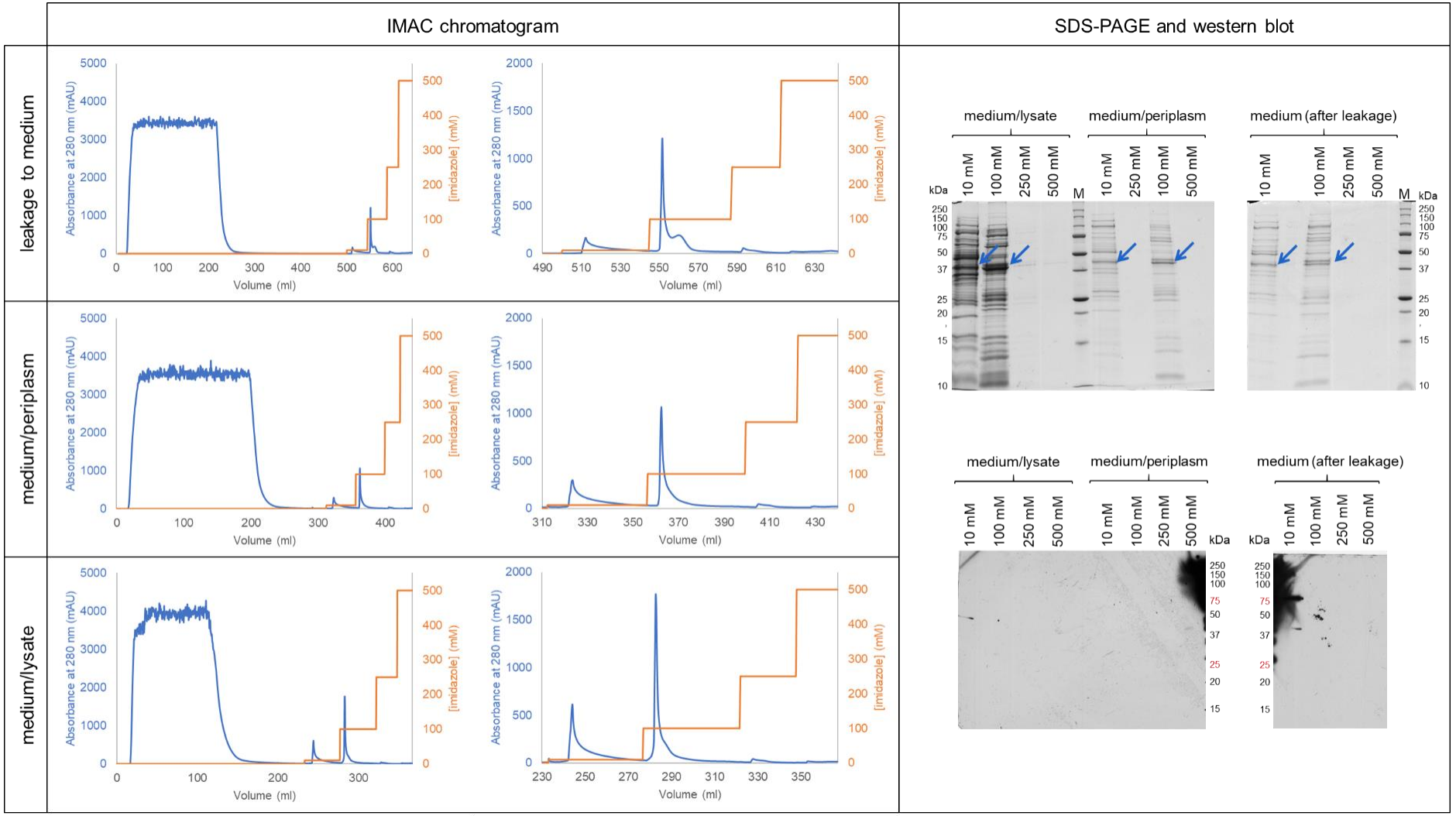

Figure 5.15: IMAC purification of FDH (43 kDa) after recombinant production in *E. coli*. The 300 ml culture was split in three. Firstly, FDH was extracted by OM permeabilization. Secondly, FDH was extracted from the periplasm by osmotic shock. Thirdly, cells were lysed. Due to medium leakage after a prolonged induction phase, medium was collected and included in all three cases. For all purifications the complete chromatogram is shown on the left. The right chromatogram shows the elution profile in more detail. All elution peaks were analysed by SDS-PAGE and western blot, shown in the right column. The blue arrows indicate the cut out gel bands, which were mainly identified to be elongation factor Tu by MALDI-TOF MS.

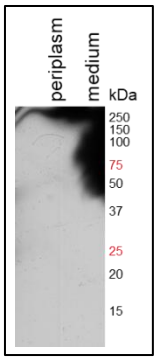

Figure 5.16: Western blot analysis of concentrated IMAC flow through samples, taken during the IMAC purification shown in Figure 5.15.

Due to these low yields, it was hypothesized that the protein might be toxic, leading to plasmid instability. Therefore, the problem of leaky expression was addressed. Firstly, the *fdh* construct was transformed in two alternative BL21 strains. The pLysS host was chosen, where basal expression is controlled by a small amount of T7 lysozyme, a natural inhibitor of T7 RNA polymerase. The second host of choice was the BL21-AI strain, with the T7 RNA polymerase under the control of the more tightly regulated *araBAD* promoter. In both strains, FDH production was obtained, as shown in Figure 5.17, and higher optical densities were measured than in case of the BL21(DE3) strain. The OD_{600} after FDH production in the pLysS host was 4.46 compared to 4.55 in the non-induced control. After expression in the BL21-AI strain, the OD_{600} was 4.16 compared to 4.85 in the non-induced control. A second approach to control basal expression, was the growth and expression of the original strain in autoinduction medium. From the moment the culture became slightly turbid, culture samples were taken for OD_{600} measurement and analysis of total cell protein (TCP) by western blot (Figure 5.18A). At the start of the exponential phase, FDH was produced. Throughout the exponential phase, no FDH was present anymore. In parallel, the *fdh* construct was retransformed in BL21(DE3), now plated out on defined medium instead of LB to make sure it lacks any inducing component. Also the preculture was set up in this defined medium instead of LB. Autoinduction starting from this preculture resulted in a slower growth and the OD_{600} after 12 h was significantly lower than inoculating autoinduction medium with an LB preculture (Figure 5.18B). On the other hand, FDH is clearly produced. These results strengthen the hypothesis that indeed FDH production is toxic for *E. coli*.

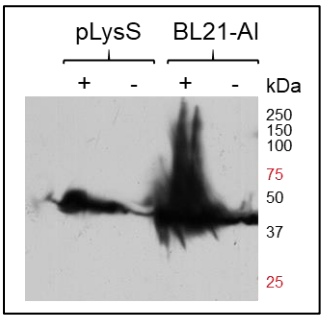

Figure 5.17: Western blot analysis of TCP samples taken after FDH (43 kDa) production in either a pLysS host or BL21-AI. In both cases, a non-induced control (-) was grown in parallel to an induced culture (+).

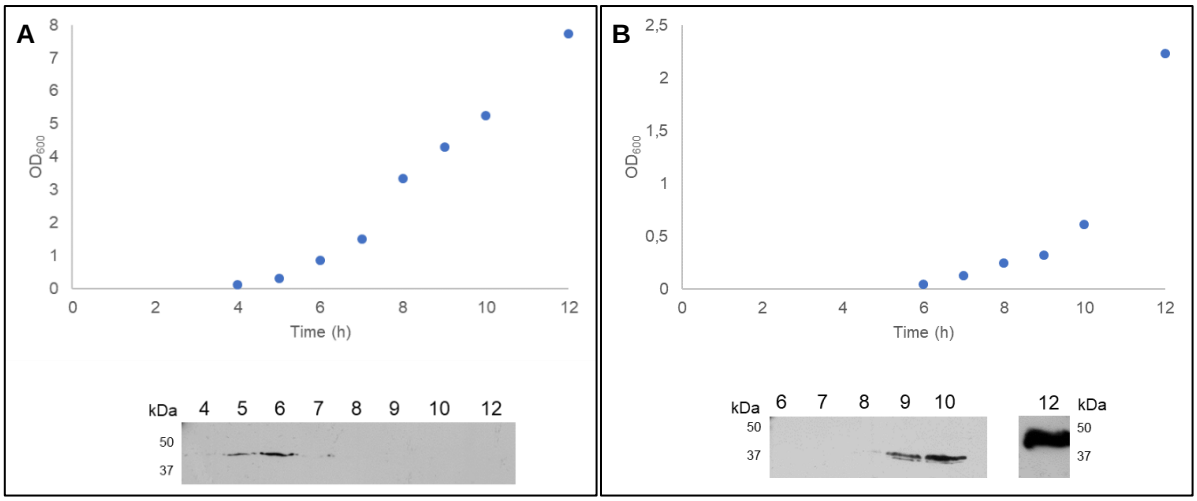

Figure 5.18: Growth curve of BL21(DE3) producing FDH (43 kDa) in autoinduction medium. A TCP sample was taken for subsequent western blot analysis, shown below the growth curve. A: LB medium was used for the agar plates and preculture. B: Non-induction medium was used for the agar plates and preculture.

4 Discussion

The production and secretion of FDH from *B. stabilis* in the recombinant production host *P. pastoris* was accomplished. However, very low yields were obtained and the enzyme degraded during the purification process. Several alternative expression conditions were tested, but did not result in significantly different results. It was seen that glycerol addition led to increased stability, as the activity was retained after a second freeze-thaw cycle compared to a decreased activity in absence of glycerol. Unfortunately, including glycerol in the buffers did not result in a higher purification yield. Because we did not observe intracellular degradation, and adapting the expression conditions in order to improve folding did not alter the final yields significantly, it was hypothesized that it was not a folding or secretion issue. This led us to believe that FDH was either synthesized to a low amount or was already degraded during the expression phase. FDH enzymes are known to have low thermostability. Could the 48 h incubation at 28 °C have resulted in the observed low yield? No significant yield increase was observed at 16 °C, thereby excluding this rationale. Alternatively, chemical instability has been reported and often was ascribed to oxidation of solvent-accessible cysteines. Indeed, when secreting FDH to the medium, the enzyme is present in an oxidizing environment, instead of in its native reducing environment of the cytoplasm. Mutation of these residues might be performed to further elucidate this hypothesis. Of note, an alignment of FDH from *B. stabilis* to the extensively investigated FDH from *C. boidinii* and *Pseudomonas* sp. 101, shows that two cysteines of the three cysteines align with Cys5 and Cys354 from *Pseudomonas* sp. 101, which have been reported to be not essential for chemical stability [375]. The third cysteine, Cys183 aligns with Cys182 from *Pseudomonas* sp. 101 and is not conserved in *C. boidinii*. It was reported that this residue was not accessible to modifying agents and was therefore not subjected to mutagenesis studies [375]. Therefore, this short analysis led to the hypothesis that cysteine modification was not the main issue here.

Two peaks are observed in the SEC chromatogram, one smaller peak eluting between 70 and 80 ml, and a broader peak eluting later. FDH enzymes are reported to be dimers, so it was hypothesized that the smaller peak coincides with the dimer. The monomer would than coincide with the broader peak and is particularly prone to degradation. Of note, Hatrongjit and Packdibamrung reported the FDH from *B. stabilis* to be an octamer instead of a dimer [382], which was not seen here. It was thus hypothesized that upon production and secretion of FDH in *P. pastoris*, the multimeric quaternary structure was lost, thereby resulting in an inherently unstable FDH protein. In order to characterize the multimeric state of FDH further, protein standards with a known MW might be run on the same SEC column in order to set up a standard curve. Alternatively, dynamic light scattering can provide further information on the multimeric structure. Two reports were found in literature on the production and secretion of FDH in *P. pastoris*. First, Takacs et al mention the recombinant production of the well-known NAD$^+$-dependent FDH *C. boidinii*, already used in industry [385]. In a second publication, the NAD$^+$-dependent FDH from *Chaetomium thermophilum*, a thermophilic fungus was successfully secreted [386]. It was hypothesized that both FDHs are more stable, as the first has been successfully used in industry and the second enzyme is derived from a thermophilic fungus. Of

note, dimerization was not discussed and Takacs et al mention that FDH from *C. boidinii* produced in *P. pastoris* showed a decreased activity compared to the enzyme purified after production in *E. coli*.

Due to stability issues in *P. pastoris*, a switch was made to the bacterial host *E. coli*. Unfortunately, FDH could not be isolated from this expression host in an active form. Although FDH was identified in a soluble form at the right MW in different cell fractions after a first expression test, no enzyme could be purified by IMAC after upscaling. The C-terminal Histag could not bind tightly to the IMAC column, unless FDH was degraded, as only a 25 kDa band appeared in the 100 mM elution peak, whereas all other signals were observed in the 10 mM imidazole washing step and/or at a higher MW in several cases. The enzyme is reduced and denatured in Laemmli buffer in preparation of SDS-PAGE. However, it was wondered if the dimer was not completely denatured, thereby giving rise to these higher MW bands. In this respect, it might be of interest to add urea in the Laemmli buffer. In the same train of thoughts, is the Histag shielded in the dimer, thereby reducing its affinity to the IMAC column? The dimeric interface is located in the central part of the polypeptide chain, so this seems unlikely. However, the periplasm is an oxidizing environment and inter subunit disulphide bridges might be formed between cysteine residues of the FDH enzyme, not present in the native dimer, thereby shielding the Histag. The results indicated that the FDH might be toxic to the cell. Especially the autoinduction assay insinuated this. Since the publication of these NADP$^+$-dependent FDHs in the BCC in 2010, no further reports were found where this enzyme was recombinantly produced and used as a regeneration system, until very recently by Jiang et al, who reported a mutation study for improved cofactor and substrate affinity [389]. They successfully produced this FDH intracellularly in *E. coli* and used a vector containing the T7*lac* promoter instead of the “plain” T7 promoter, resulting in a more controlled basal expression. However, the authors do not report actual protein yields and they refer also on difficulties on expression and purifications, without details.

5 Conclusion

The FDH from *B. stabilis* is a difficult to produce protein in both *P. pastoris* and *E. coli*. When secreted in the yeast *P. pastoris*, the problem appeared to be in the fact that no multimeric complex was formed, whereas periplasmic localization and/or protein toxicity seemed to be a problem in *E. coli*. Overall, the recombinant production of FDH is accompanied with many issues and more efforts are required to resolve these issues. Our efforts on the production of FDH for cofactor regeneration was only relevant when an active self-sufficient CYP is obtained for DCA production. Therefore, we focused on the development of such a CYP and the work on FDH was put on hold.

Chapter 6
General conclusion and future perspectives

In this dissertation, the development of an *in vitro* cell-free biocatalytic process for the production of long-chain α,ω-dicarboxylic acids (DCAs) was pursued. Several arguments favoured the choice for an *in vitro* process. The use of isolated enzymes circumvents the issue of undesired side reactions such as the degradation of substrate and product in the β-oxidation pathway. The substrate and product do not need to cross any membrane barrier. This is in contrast to whole-cell bioconversion, where the β-oxidation needs to be inhibited or knocked out and membrane transport is a major limitation in long-chain DCA production. Furthermore, the implementation of an *in vitro* process in the classical chemical industry is more straightforward as fermentation processes require expertise and specialized equipment, demanding huge investments in this sector. Several CYPs have been described to selectively hydroxylate the ω-position of fatty acids and overoxidation to DCAs has been reported as well. One of these described CYPs, i.e. CYP52A13 from the industrial DCA producer *Candida viswanathii* [23], [24], was selected for recombinant production in search for a suitable biocatalyst. In the first experimental chapter (chapter 3), it was investigated whether it was feasible to produce a soluble and self-sufficient variant of CYP52A13 in the recombinant host *Pichia pastoris*. A soluble enzyme was preferred in order to obtain a flexible system, easy to implement in industry. In chapter 3, we illustrate that production of a soluble CYP52A13 was accomplished, as well as the production of its reductase partner CPR-a in a soluble form. Even more, both enzymes were obtained in a secreted form, although activity could only be shown for the reductase. In both cases, the N-terminal tag, described by Lamb et al for the crystallization of CPR from *Saccharomyces cerevisiae* [193], was used. When both proteins were fused in order to create a chimer for the construction of a self-sufficient CYP system, it was seen that upon washing the microsome with a high salt buffer, the chimer appeared in the soluble supernatant. Furthermore, replacement of the N-terminal membrane anchor with this artificial N-terminal tag seemed to improve the enzyme yield. This N-tag thus proves to be a valuable alternative sequence for recombinant class II CYP production.

The CYP monooxygenase superfamily contains many potential members for industrial application, exemplified by the successful CYP application in the fermentative production of artemisinic acid, the precursor of the antimalarial drug artemisin [132]. Unfortunately, they have some disadvantages that need to be overcome, such as a low specific activity, but also instability and the need for a redox partner and stoichiometric amounts of NAD(P)H limits their use [120], [122]. The class II type CYPs have the additional challenge of being membrane-bound proteins [134]. Due to aforementioned disadvantages, current CYP applications are constraint to whole-cell bioconversions and creating an *in vitro* cell-free biocatalytic process is not straightforward, requiring a lot of research. The troublesome production of CYP52A13 in *P. pastoris* demonstrated the complexity of this enzyme superfamily and many different factors determine the success or failure of this enzyme system. Which truncations in the membrane binding domains of the CYP enzyme results in the highest protein yield in a recombinant host without affecting the activity? Which redox partner results in the highest coupling efficiency, without negatively influencing the CYP activity? How is the N-terminus of the redox partner best modified to allow for the most efficient electron transfer and which linker supports this reduction the best? Which recombinant host is suited the most? To add to this complexity, the

involvement of cytochrome *b5* needs to be addressed as well as the influence of the membrane in the overall catalysis To obtain the best performing self-sufficient biocatalysts, many parameters need to be considered and a high-throughput screening method would therefore be of great value to further investigate this. For example, the proposed high-throughput screening method from Gudiminchi, where CYP expression was determined by measuring the CO-bound difference spectrum of whole-cells, is an interesting approach [327]. This might possibly be extended to a method where also the reductase partner and the coupling efficiency are evaluated, e.g. by including the assay proposed by Hayashi et al where measurement of the CO-bound difference spectrum was used to prove the ability of the redox partner to reduce the heme group [341].

During this dissertation, many issues were encountered in the development of a soluble class II CYP. Based on the stability issues and absence of an active system, we dare to conclude that the implementation of a solubilized class II CYP as an industrial biocatalyst is currently not feasible. Even if alternative truncations and chimers would lead to an active enzyme, the lack of stability would limit the biocatalytic performance and the recycling of the biocatalyst would not be possible. Even though the enzymes, naturally oxidizing the ω-position, show a high regioselectivity and the final purity of the formed DCA might be increased compared to fermentative production, thereby reducing costs regarding DCA purification, this would not compensate for the costs of the biocatalyst. Thereby, no cost-effective alternative could be delivered using a soluble self-sufficient enzyme. Alternatively, the soluble CYP153 family of class I has been reported to regioselectively oxidize the ω-position [82], [88], which might provide better candidates for the development of a soluble self-sufficient biocatalyst. However, the use of these enzymes did not yet lead to conversion yields in the range of *C. viswanathii* and overoxidation activity is almost nonexisting. The soluble self-sufficient CYP102A1 of class VIII shows the highest specific activity reported to date and hydroxylates FAs, although preferably at the positions ω-1, ω-2 and ω-3. Therefore, this is an interesting candidate for *in vitro* DCA production. Mutation studies have been reported and ω-oxidation was accomplished. However, the selectivity of the natural ω-oxidizing CYPs was superior. The mutated CYP102A1 showed a regioselectivity of 74 % whereas natural ω-oxidizing CYPs show selectivities of over 90 % [100]. Furthermore, limited stability has also been reported for soluble CYPs [120]. In this regard, the characterization of an increasing number of thermophilic CYPs raises interest [141]. Unfortunately, no FA hydroxylating thermophilic CYP has been identified to date.

Even though no biocatalyst was obtained, the artificial N-terminal tag showed to solubilize CYP and the chimer to some extent and seemed to increase enzyme yield. Therefore, this tag might serve as an approach to produce soluble protein for other purposes, such as for crystallization efforts. No structure of a CYP52 family member has been resolved thus far. Recently, the crystal structures of other ω-hydroxylating CYPs have been determined. For example, Hsu et al determined the structure of the rabbit CYP4B1 in a substrate (octane) bound form [390]. Here, it was described how the Ala or Gly from the consensus sequence (A/G)GX(D/E)T on helix I was substituted by Glu. An ester bond between this Glu and the 5-methyl group of heme

resulted in the architecture favouring ω-hydroxylation over subterminal hydroxylation. However, the CYP4 family is the only CYP family reported to covalently bind the heme group, and although homology modelling of CYP52A23 to CYP4B1 gave some structure-function relationship insights [53], structures of the CYP52 family would be of value. Hoffmann et al resolved the structure of the soluble ω-hydroxylating $CYP153_{M.aq.}$ in a substrate-free form as well as with the substrate dodecanoic acid bound [391]. They demonstrated the benefit of having the crystal structure at hand. Guided by the available structure, targeted mutational studies were performed, resulting in a variant with a lower K_M by introducing a so-called substrate anchor in the active site. Moreover, the flexibility of the substrate recognition site was increased, further improving the activity towards dodecanoic acid [391]. Indeed, crystal structures have led to the rational and semi-rational design of several CYPs, targeting not only the substrate recognition sites, but also the substrate access channels and the CPR interaction site, significantly improving overall activity [256], [392]. This is of importance due to the fact that CYPs are generally known to have a low specific activity [120], [122]. Furthermore, engineering a better substrate fit might positively influence the coupling efficiency [176]. Additionally, it might be of interest to learn how the regioselectivity is obtained. CYPs with characteristics more suited for industrial application, such as CYP102A1 or thermostable CYPs, might be engineered by rational design in order to introduce the same high regioselectivity.

As the structure of the highly similar CPR-b from *C. tropicalis* is available [195], as well as the structure of the CYP102A1 heme-FMN complex [190], docking studies of CPR-b with the CYP52A13 structure, guided by the heme-FMN domain structure of CYP102A1, might identify potential reductase interaction sites and could be targeted to further improve coupling efficiency and thereby enhancing overall activity. Additional studies such as mutation studies and cryo-EM can lead to further insights in this complex formation. It has been reported that the reductase partner can have a marked influence on the CYP activity. Does the CPR redox partner aids in the regioselectivity of CYP52A13? This information as well could be used for the enhancement of regioselectivity in other CYPs.

An interesting approach to combine advantages of both whole-cell bioconversions and *in vitro* biocatalysis, is the surface display of CYP enzymes on the outer membrane (OM) of *Escherichia coli*. On the one hand, stability of the produced CYP is enhanced by *in vivo* immobilization into a membrane environment and no expensive enzyme purification process is required, advantages accompanied with whole-cell bioconversions. On the other hand, no membrane barrier needs to be crossed, a big advantage accompanied with *in vitro* biocatalysis [252]. In literature, three different surface display anchors were shown to integrate the self-sufficient CYP102A1 in the OM of *E. coli*, i.e. a truncated ice-nucleation protein [250], the autotransporter β-barrel (Autodisplay) [254] and the SpyCatcher-SpyTag system [255]. The Autodisplay approach was investigated in more detail and enabled a total turnover number of 54 700, the highest reported for CYP enzymes [254]. Autodisplay was found to be applicable for the surface display of mammalian class II CYPs as well [247], [249]. In this dissertation, it was found that the use of a truncated ice-nucleation protein InaK could be extended to the surface display of self-sufficient class II CYPs, exemplified in chapter 4 for CYP52A13, fused

to the reductase domain of CYP102A1. Unfortunately, no oleic-acid converting system was obtained. Even though no active system was created, it is still believed to be worth considering for other class II CYPs, as CYP102A1 has proven to be a valuable model enzyme for class II CYPs. CYP52A13 proved to be difficult to produce in *P. pastoris* so it cannot be excluded that the lack of activity can be ascribed to inherent problems with the enzyme rather than the surface display system.

In chapter 4, it was investigated if this surface display approach could be taken one step further, i.e. displaying a self-sufficient CYP on outer membrane vesicles (OMVs) derived from the *E. coli* OM. Park et al showed how OMV display led to an increased activity compared to yeast surface display of the same enzyme complex [349]. It was hypothesized that the increased activity could be ascribed to the nanoscale dimensions of the OMVs, resulting in an increased enzyme:volume ratio and improved substrate accessibility [349], [350]. In this dissertation, membrane vesicles were successfully isolated, either EDTA-extracted eOMVs or spontaneously formed sOMVs (which might include vesicles as a result of cell lysis) and showed to display CYP102A1, actively converting the model substrate 12-pNCA. The four different alternative CYP52A13 chimers were also found in isolated membrane vesicles. OMVs have proven their value in the development of a vaccine against *Neisseria meningitis* and are increasingly investigated as a vaccine against other pathogens [348]. Furthermore, the use of engineered OMVs has emerged as vehicles for targeted drug delivery [344]. In this chapter, it was shown that OMVs can serve as an *in vivo* immobilization strategy for recombinant production of the challenging class II CYP monooxygenases, thereby proposing an alternative production strategy.

The OMV display approach shows great promise, although there are still many parameters that need to be investigated and optimized before this would result in an economically attractive process. Firstly, it is of interest to obtain actual sOMVs, obtained from budding of the OM. This would allow for a continuous process where the medium is removed for sOMV isolation instead of a batch process where the cells need to be collected for eOMV extraction or due to cell lysis. In this context, other hypervesiculating mutants might be investigated as well, e.g. mutants with disruptions of genes involved in the σ^E stress response pathway [356]. Secondly, the stability of these OMVs needs to be assessed. OMVs have been reported to be stable during long-term storage and upon several cycles of freeze-thawing [393], [394]. However, the stability in the desired biocatalytic setup where e.g. shearing forces and solvent will be present, needs to be tested. Thirdly, it might be of interest to engineer *E. coli* for the inclusion of a heme biosynthetic pathway. The heme precursor δ-aminolevulinic acid (5-ALA) is expensive so elimination of this additive would greatly reduce the production costs. For example, Harnastai et al showed that co-expression of glutamyl-tRNA reductase proved to be sufficient for eliminating the need for 5-ALA supplementation [239]. Another parameter that needs attention is the purification of these vesicles. OMVs are now generally collected by ultracentrifugation or ultrafiltration techniques. Alternatively, it has been shown that IMAC can be used for the purification of OMVs [395], thereby only retaining vesicles displaying the protein of interest. On an industrial scale, both ultrafiltration and IMAC are established but expensive purification processes.

Therefore, it remains to be seen if OMVs would be a suitable nanobiocatalyst for the production of bulk chemicals. It should be noted that new strategies for the recombinant production of CYPs are valuable in synthetic biology applications, not only for bulk chemicals such as long-chain DCAs, but in the synthesis of fine chemicals as well. For example in the pharmaceutical industry, CYPs are involved in the production of many interesting compounds (see Chapter 1 - 2.1). Moreover, liver CYP enzymes are of major importance in drug metabolism and great interest continues to go to recombinant production systems as biosensors in drug development. Even though the applicability of OMVs in context of bulk chemical production might be questioned, it can still prove to be an invaluable immobilization strategy for other applications.

We believe that the strategy of *in vivo* immobilization in a membrane environment might pave the way to developing a biocatalyst to be used in the chemical industry, given the difficulties associated with the production of a soluble CYP and OMV display might thus serve as a valuable tool. Alternatively, other *in vivo* immobilization strategies have been proposed, such as the integration of CYPs in the membrane of polyhydroxybutyrate granules [306]. Another interesting idea was found for the recombinant production of rat CPR. Yim et al reported recombinant expression in the host *Bacillus subtilis* and *in vivo* immobilization of the enzyme by display on the surface of spores [396]. These might be worth investigating in parallel in order to obtain the best possible biocatalyst.

A matter that has shortly been addressed in chapter 5 concerns cofactor regeneration in order to make the biocatalytic process economically feasible. Formate dehydrogenase (FDH) is recognized as the ideal regeneration system for industrial application due to its many advantages. The substrate formate is cheap, readily available and only one enzyme was reported to be inhibited by formate so far. The reaction is irreversible and the formed product CO_2 is easily removed, thereby not interfering with downstream processing of the product of interest. Unfortunately, some disadvantages have limited the use of FDH in industry, such as low chemical and thermal stability and the specificity for NAD^+. Although many protein engineering efforts have led to significant improvements, the specificity for NAD^+ remains an issue up to today. Three FDH enzymes have been reported in literature to accept $NADP^+$. Two of them prefer acidic conditions [383], [384], whereas the FDH from *Burkholderia stabilis* shows a broader pH range [382]. However, this last FDH proved to be difficult to produce recombinantly, as shown in chapter 5. The OMV display approach as discussed in chapter 4, might actually serve as an immobilization strategy for this FDH as well. The FDH from *C. boidinii* has successfully been surface displayed using a INP anchor [387]. The proposed strategy in chapter 4 might therefore be extended to FDH display on the surface of OMVs. In order to create an *in vitro* biocatalyst which is truly self-sufficient, and thus also including NADPH regeneration, it has been proposed in literature to co-immobilize the regenerating enzyme with the self-sufficient CYP. For example, the self-sufficient CYP102A1 has been co-immobilized with an engineered FDH in a sol-gel matrix [293] and with GDH on an anionic sulfopropyl-activated carrier (ReliSorb SP) [298]. In the same train of thoughts, co-expression and OMV display of both the self-sufficient CYP and a regeneration system might be investigated to create such a true self-sufficient system. Furthermore, co-expression would

allow for an approach where all enzymes are isolated in one and the same purification process. A fusion construct between CYP102A1 and PTDH was created as well in the search for a true self-sufficient enzyme [294]. Would it be possible to display this fusion on the surface of *E. coli*, and by extension, on OMVs?

In the context of cofactor regeneration, it has been reported for the self-sufficient CYP102A1 to switch the cofactor specificity from NADPH to NADH [296]. Alternatively, it might be advantageous to test different heterologous redox partners and extend the search to redox partners from other classes which show NADH specificity. Indeed care should be taken when changing the redox partner and this can negatively influence the conversion activity, substrate selectivity and coupling efficiency. However, in some cases switching the redox partner positively influenced aforementioned parameters [58], [264], [392]. As cofactor recycling is an integral part of the biocatalytic process, the NADH specificity might result in decreased costs, as NADH is cheaper, compensating for possible issues caused by switching to a heterologous redox partner. When screening for the best redox partner, it is thus important to also take this into account and to consider the overall process.

In conclusion, the CYP superfamily is a very interesting yet challenging group of enzymes for industrial application. Their use is now limited and whole-cell bioconversions are preferred. In the production of long-chain DCAs, CYPs have proven of value. The fermentative production of long-chain DCAs with the yeast *C. viswanathii*, where CYPs are responsible for the first and rate-limiting step of the ω-oxidation pathway, results in good yields and circumvents many disadvantages of the chemical conversion of FAs to DCAs by means of metathesis. However, the fermentation process is executed with an opportunistic pathogen and delivers only a limited set of DCAs at elevated prices. Therefore, an *in vitro* cell-free approach was investigated in this dissertation, aiming to obtain a flexible system resulting in high productivity and purity, ultimately leading to a wide variety of long-chain DCAs at decreased costs. However, no active soluble self-sufficient CYP could be created. Based on the obtained results, it was concluded that the use of class II CYPs as a soluble biocatalyst is currently not feasible. Both the production process of the unstable biocatalysts as well as the poor performance, would lead to elevated costs, consequently elevating the costs of the final DCA product. Immobilization is a well-known strategy to increase stability of the biocatalyst. In order to circumvent a costly purification before continuing to immobilization, an *in vivo* immobilization strategy was proposed, targeting CYPs to the surface of OMVs. A proof-of-concept was delivered using the model enzyme CYP102A1, but much more research is required in order to assess its feasibility as a vehicle for industrial application. Although no industrial biocatalyst could be delivered for the production of long-chain α,ω-DCAs, this dissertation provided tools with great promise in other applications, e.g. in the development of biosensors.

References

[1] S. Huf, S. Krügener, T. Hirth, S. Rupp, and S. Zibek, “Biotechnological synthesis of long-chain dicarboxylic acids as building blocks for polymers,” *Eur. J. Lipid Sci. Technol.*, vol. 113, pp. 548–561, 2011.
[2] H. Lee, Y. E. C. Sugiharto, H. Lee, W. Jeon, J. Ahn, and H. Lee, “Biotransformation of dicarboxylic acids from vegetable oil–derived sources: current methods and suggestions for improvement,” *Appl. Microbiol. Biotechnol.*, vol. 103, pp. 1545–1555, 2019.
[3] N. Werner and S. Zibek, “Biotechnological production of bio-based long-chain dicarboxylic acids with oleogenious yeasts,” *World J. Microbiol. Biotechnol.*, vol. 33, no. 194, pp. 1–9, 2017.
[4] K. Kroha, “Industrial biotechnology provides opportunities for commercial production of new long-chain dibasic acids,” *Inform*, vol. 15, no. 9, pp. 568–571, 2004.
[5] K. D. Green, M. K. Turner, and J. M. Woodley, “*Candida cloacae* oxidation of long-chain fatty acids to dioic acids,” *Enzyme Microb. Technol.*, vol. 27, pp. 205–211, 2000.
[6] E. J. Vandamme and W. Soetaert, “Bioflavours and fragrances via fermentation and biocatalysis,” *J. Chem. Technol. Biotechnol.*, vol. 77, pp. 1323–1332, 2002.
[7] F. Stempfle, P. Ortmann, and S. Mecking, “Long-chain aliphatic polymers to bridge the gap between semicrystalline polyolefins and traditional polycondensates,” *Chem. Rev.*, vol. 116, pp. 4597–4641, 2016.
[8] “Long Chain Dicarboxylic Acid Market Size | Industry Report, 2018-2025.” [Online]. Available: https://www.grandviewresearch.com/industry-analysis/long-chain-dicarboxylic-acid-market. [Accessed: 05-Mar-2019].
[9] C. Jiménez-Rodriguez, G. R. Eastham, and D. J. Cole-Hamilton, “Dicarboxylic acid esters from the carbonylation of unsaturated esters under mild conditions,” *Inorg. Chem. Commun.*, vol. 8, pp. 878–881, 2005.
[10] P. B. van Dam, M. C. Mittelmeijer, and C. Boelhouwer, “Metathesis of unsaturated fatty acid esters by a homogeneous tungsten hexachloride–tetramethyltin catalyst,” *J. Chem. Soc., Chem. Commun.*, no. 22, pp. 1221–1222, 1972.
[11] T. M. Trnka and R. H. Grubbs, “The development of L2X2RU=CHR olefin metathesis catalysts: An organometallic success story,” *Acc. Chem. Res.*, vol. 34, no. 1, pp. 18–29, 2001.
[12] G. C. Vougioukalakis and R. H. Grubbs, “Ruthenium-based heterocyclic carbene-coordinated olefin metathesis catalysts,” *Chem. Rev.*, vol. 110, no. 3, pp. 1746–1787, 2010.
[13] H. L. Ngo, K. Jones, and T. A. Foglia, “Metathesis of unsaturated fatty acids: Synthesis of long-chain unsaturated-α,ω-dicarboxylic acids,” *J Am Oil Chem Soc*, vol. 83, no. 7, pp. 629–634, 2006.
[14] H. L. Ngo and T. A. Foglia, “Synthesis of long chain unsaturated-α,ω-dicarboxylic acids from renewable materials via olefin metathesis,” *J. Am. Oil Chem. Soc.*, vol. 84, pp. 777–784, 2007.
[15] S. Chikkali and S. Mecking, “Refining of plant oils to chemicals by olefin metathesis,” *Angew. Chemie - Int. Ed.*, vol. 51, pp. 5802–5808, 2012.
[16] H. Clavier, K. Grela, A. Kirschning, M. Mauduit, and S. P. Nolan, “Sustainable concepts in olefin metathesis,” *Angew. Chemie - Int. Ed.*, vol. 46, no. 36, pp. 6786–6801, 2007.
[17] S. Hübner, J. G. De Vries, and V. Farina, “Why does industry not use immobilized transition metal complexes as catalysts?,” *Adv. Synth. Catal.*, vol. 358, no. 1, pp. 3–25, 2016.
[18] “Elevance Inherent® C18 Diacid.” [Online]. Available: https://elevance.com/product/elevance-inherent-c18-diacid/?prod-id=426&cat-id=polymer_building_blocks-13#0. [Accessed: 21-Jan-2021].
[19] T. E. Snead, S. Cohen, and D. L. Gildon, “Methods of refining and producing dibasic

esters and acids from natural oil feedstocks," US 9,284,512 B2, 2016.
[20] S. Cohen, M. l. Luetkens, C. Balakrishnan, and R. Snyder, "Methods of refining and producing fuel from natural oil feedstocks," WO 2011/046872 A2, 2011.
[21] M. Arie, H. Matsuda, K. Furuhashi, and M. Takagi, "Phylogenetic identification of n-alkane assimilating *Candida* yeasts based on nucleotide divergence in the 5′ end of LSU rDNA gene," *J. Gen. Appl. Microbiol.*, vol. 46, pp. 257–262, 2000.
[22] S. A. Shankarnarayan *et al.*, "Molecular typing and antifungal susceptibility of *Candida viswanathii*, India," *Emerg. Infect. Dis.*, vol. 24, no. 10, pp. 1956–1958, 2018.
[23] D. L. Craft, K. M. Madduri, M. Eshoo, and C. R. Wilson, "Identification and characterization of the CYP52 family of *Candida tropicalis* ATCC 20336, important for the conversion of fatty acids and alkanes to α,ω-dicarboxylic acids," *Appl. Environ. Microbiol.*, vol. 69, no. 10, pp. 5983–5991, 2003.
[24] W. H. Eschenfeldt *et al.*, "Transformation of fatty acids catalyzed by cytochrome P450 monooxygenase enzymes of *Candida tropicalis* transformation of fatty acids catalyzed by cytochrome P450 monooxygenase enzymes of *Candida tropicalis*," *Appl. Environ. Microbiol.*, vol. 69, no. 10, pp. 5992–5999, 2003.
[25] S. Picataggio, K. Deanda, and J. Mielenz, "Determination of *Candida tropicalis* acyl coenzyme A oxidase isozyme function by sequential gene disruption.," *Mol. Cell. Biol.*, vol. 11, no. 9, pp. 4333–4339, 1991.
[26] S. Picataggio *et al.*, "Metabolic engineering of *Candida tropicalis* for the production of long–chain dicarboxylic acids," *Nat. Biotechnol.*, vol. 10, pp. 894–898, 1992.
[27] D. Eirich *et al.*, "Cloning and characterization of three fatty alcohol oxidase genes from *Candida tropicalis* strain ATCC 20336," *Appl. Environ. Microbiol.*, vol. 70, no. 8, pp. 4872–4879, 2004.
[28] Z. Cao, H. Gao, M. Liu, and P. Jiao, "Engineering the acetyl-CoA transportation system of *Candida tropicalis* enhances the production of dicarboxylic acid," *Biotechnol. J.*, vol. 1, pp. 68–74, 2006.
[29] P. Mishra *et al.*, "Genome-scale metabolic modeling and in silico analysis of lipid accumulating yeast *Candida tropicalis* for dicarboxylic acid production.," *Biotechnol. Bioeng.*, vol. 113, no. 9, pp. 1993–2004, 2016.
[30] Z. H. Yi and H. J. Rehm, "Metabolic formation of dodecanedioic acid from n-dodecane by a mutant of *Candida tropicalis*," *Eur. J. Appl. Microbiol. Biotechnol.*, vol. 14, pp. 254–258, 1982.
[31] D. Fabritius, H. J. Schäfer, and A. Steinbüchel, "Identification and production of 3-hydroxy-Δ9-cis-1,18-octadecenedioic acid by mutants of *Candida tropicalis*," *Appl. Microbiol. Biotechnol.*, vol. 45, pp. 342–348, 1996.
[32] S. Gangopadhyay, S. Nandi, and S. Ghosh, "Biooxidation of fatty acid distillates to dibasic acids by a mutant of *Candida tropicalis*.," *J. Oleo Sci.*, vol. 56, no. 1, pp. 13–17, 2006.
[33] I. Shiio and R. Uchio, "Microbial production of long-chain dicarboxylic acids from n-alkanes. Part I.," *Agric. Biol. Chem.*, vol. 35, no. 13, pp. 2033–2042, 1971.
[34] S. Liu, C. Li, X. Fang, and Z. Cao, "Optimal pH control strategy for high-level production of long-chain α,ω-dicarboxylic acid by *Candida tropicalis*," *Enzyme Microb. Technol.*, vol. 34, pp. 73–77, 2004.
[35] P. Jiao, Y. Huang, S. Li, Y. Hua, and Z. U. A. Cao, "Effects and mechanisms of H_2O_2 on production of dicarboxylic acid," *Biotechnol. Bioeng.*, vol. 75, no. 4, pp. 456–462, 2001.
[36] I. Funk, N. Rimmel, C. Schorsch, V. Sieber, and J. Schmid, "Production of dodecanedioic acid via biotransformation of low cost plant-oil derivatives using *Candida tropicalis*," *J. Ind. Microbiol. Biotechnol.*, vol. 44, pp. 1491–1502, 2017.

[37] Y. E. C. Sugiharto *et al.*, "Effect of decanoic acid and 10-hydroxydecanoic acid on the biotransformation of methyl decanoate to sebacic acid," *AMB Express*, vol. 8, no. 75, pp. 1–9, 2018.
[38] I. Funk, V. Sieber, and J. Schmid, "Effects of glucose concentration on 1,18-cis-octadec-9-enedioic acid biotransformation efficiency and lipid body formation in *Candida tropicalis*," *Sci. Rep.*, vol. 7, no. 13842, pp. 1–11, 2017.
[39] J. Bauwelinck, M. Wijnants, S. Tavernier, and I. Cornet, "The evaluation of oleic acid alternatives for the biochemical production of 9-octadecenedioic acid," *Biochem. Eng. J.*, vol. 161, no. 107660, 2020.
[40] W. Cao, H. Li, J. Luo, J. Yin, and Y. Wan, "High-level productivity of α,ω-dodecanedioic acid with a newly isolated *Candida viswanathii* strain," *J. Ind. Microbiol. Biotechnol.*, vol. 44, pp. 1191–1202, 2017.
[41] W. Cao, B. Liu, J. Luo, J. Yin, and Y. Wan, "A, Ω-Dodecanedioic acid production by *Candida viswanathii* ipe-1 with co-utilization of wheat straw hydrolysates and n-dodecane," *Bioresour. Technol.*, vol. 243, pp. 179–187, 2017.
[42] S. Picataggio, T. Rohrer, and L. D. Eirich, "Method for increasing the omega-hydroxylase activity in *Candida tropicalis*," US 5,620,878, 1997.
[43] M. Ohkuma, S. I. Muraoka, T. Tanimoto, M. Fujii, A. Ohta, and M. Takagi, "CYP52 (Cytochrome P450alk) multigene family in *Candida maltosa*: Identification and characterization of eight members," *DNA Cell Biol.*, vol. 14, no. 2, pp. 163–173, 1995.
[44] T. Zimmer, M. Ohkuma, A. Ohta, M. Takagi, and W. H. Schunck, "The CYP52 multigene family of *Candida maltosa* encodes functionally diverse n-alkane-inducible cytochromes p450," *Biochem. Biophys. Res. Commun.*, vol. 224, no. 3, pp. 784–789, 1996.
[45] D. R. Nelson, "Cytochrome P450 nomenclature, 2004.," in *Methods in Molecular Biology*, 2nd ed., vol. 320, I. R. Phillips and E. A. Shephard, Eds. Humana Press Inc., Totowa, NJ, 2006, pp. 1–10.
[46] U. Scheuer, T. Zimmer, D. Becher, F. Schauer, and W. H. Schunck, "Oxygenation cascade in conversion of n-alkanes to α,ω-dioic acids catalyzed by cytochrome P450 52A3," *J. Biol. Chem.*, vol. 273, no. 49, pp. 32528–32534, 1998.
[47] A. Hara *et al.*, "Repression of fatty-acyl-CoA oxidase-encoding gene expression is not necessarily a determinant of high-level production of dicarboxylic acids in industrial dicarboxylic-acid-producing *Candida tropicalis*," *Appl. Microbiol. Biotechnol.*, vol. 56, pp. 478–485, 2001.
[48] T. Kogure, H. Horiuchi, H. Matsuda, M. Arie, M. Takagi, and A. Ohta, "Enhanced induction of cytochromes P450alk that oxidize methyl-ends of n-alkanes and fatty acids in the long-chain dicarboxylic acid-hyperproducing mutant of *Candida maltosa*," *FEMS Microbiol. Lett.*, vol. 271, pp. 106–111, 2007.
[49] R. Uchio and I. Shiio, "Microbial production of long-chain dicarboxylic acids from n-alkanes. Part II.," *Agric. Biol. Chem.*, vol. 36, no. 3, pp. 426–433, 1972.
[50] N. Werner, M. Dreyer, W. Wagner, N. Papon, S. Rupp, and S. Zibek, "*Candida guilliermondii* as a potential biocatalyst for the production of long-chain α,ω-dicarboxylic acids," *Biotechnol. Lett.*, vol. 39, pp. 429–438, 2017.
[51] H. Lee *et al.*, "Characterization of the newly isolated ω-oxidizing yeast *Candida sorbophila* DS02 and its potential applications in long-chain dicarboxylic acid production," *Appl. Microbiol. Biotechnol.*, vol. 101, pp. 6333–6342, 2017.
[52] H. Lee *et al.*, "Development of a promising microbial platform for the production of dicarboxylic acids from biorenewable resources," *Biotechnol. Biofuels*, vol. 11, no. 310, pp. 1–14, 2018.
[53] H. G. Park *et al.*, "CYP52A23 from *Candida albicans* and its substrate preference for

fatty Acid hydroxylation," *Arch. Biochem. Biophys.*, vol. 671, pp. 27–34, 2019.
[54] M. De Graeve, S. L. De Maeseneire, S. L. K. W. Roelants, and W. Soetaert, "*Starmerella bombicola*, an industrially relevant, yet fundamentally underexplored yeast," *FEMS Yeast Res.*, vol. 18, no. 7, pp. 1–13, 2018.
[55] K. Lottermoser, W. H. Schunck, and O. Asperger, "Cytochromes P450 of the sophorose lipid-producing yeast *Candida apicola*: Heterogeneity and polymerase chain reaction-mediated cloning of two genes," *Yeast*, vol. 12, pp. 565–575, 1996.
[56] I. N. A. Van Bogaert, M. Demey, D. Develter, W. Soetaert, and E. J. Vandamme, "Importance of the cytochrome P450 monooxygenase CYP52 family for the sophorolipid-producing yeast *Candida bombicola*," *FEMS Yeast Res.*, vol. 9, pp. 87–94, 2009.
[57] F. C. Huang, A. Peter, and W. Schwab, "Expression and characterization of CYP52 genes involved in the biosynthesis of sophorolipid and alkane metabolism from *Starmerella bombicola*," *Appl. Environ. Microbiol.*, vol. 80, no. 2, pp. 766–776, 2014.
[58] J. Liu, C. Zhang, and W. Lu, "Biosynthesis of long-chain ω-hydroxy fatty acids by engineered *Saccharomyces cerevisiae*," *J. Agric. Food Chem.*, vol. 67, pp. 4545–4552, 2019.
[59] R. Ledesma-Amaro and J. M. Nicaud, "*Yarrowia lipolytica* as a biotechnological chassis to produce usual and unusual fatty acids," *Prog. Lipid Res.*, vol. 61, pp. 40–50, 2016.
[60] H. Takai, R. Iwama, S. Kobayashi, H. Horiuchi, R. Fukuda, and A. Ohta, "Construction and characterization of a *Yarrowia lipolytica* mutant lacking genes encoding cytochromes P450 subfamily 52," *Fungal Genet. Biol.*, vol. 49, pp. 58–64, 2012.
[61] R. Iwama, S. Kobayashi, C. Ishimaru, A. Ohta, H. Horiuchi, and R. Fukuda, "Functional roles and substrate specificities of twelve cytochromes P450 belonging to CYP52 family in n-alkane assimilating yeast *Yarrowia lipolytica*," *Fungal Genet. Biol.*, vol. 91, pp. 43–54, 2016.
[62] M. S. Smit, M. M. Mokgoro, E. Setati, and J. M. Nicaud, "α,ω-Dicarboxylic acid accumulation by acyl-CoA oxidase deficient mutants of *Yarrowia lipolytica*," *Biotechnol. Lett.*, vol. 27, pp. 859–864, 2005.
[63] M. Gatter *et al.*, "A newly identified fatty alcohol oxidase gene is mainly responsible for the oxidation of long-chain ω-hydroxy fatty acids in *Yarrowia lipolytica*," *FEMS Yeast Res.*, vol. 14, pp. 858–872, 2014.
[64] A. Abghari, C. Madzak, and S. Chen, "Combinatorial engineering of *Yarrowia lipolytica* as a promising cell biorefinery platform for the de novo production of multi-purpose long chain dicarboxylic acids," *Fermentation*, vol. 3, no. 40, pp. 1–30, 2017.
[65] P. Mishra, N. R. Lee, M. Lakshmanan, M. Kim, B. G. Kim, and D. Y. Lee, "Genome-scale model-driven strain design for dicarboxylic acid production in *Yarrowia lipolytica*," *BMC Syst. Biol.*, vol. 12, no. Suppl 2, pp. 9–20, 2018.
[66] L. Lin, W. Fang, X. Liao, F. Wang, D. Wei, and R. J. St. Leger, "The MrCYP52 cytochrome P450 monooxygenase gene of *Metarhizium robertsii* is important for utilizing insect epicuticular hydrocarbons," *PLoS One*, vol. 6, no. 12, p. e28984, 2011.
[67] S. Zhang *et al.*, "CYP52X1, representing new cytochrome P450 subfamily, displays fatty acid hydroxylase activity and contributes to virulence and growth on insect cuticular substrates in entomopathogenic fungus *Beauveria bassiana*," *J. Biol. Chem.*, vol. 287, no. 16, pp. 13477–13486, 2012.
[68] C. Huarte-Bonnet *et al.*, "Insights into hydrocarbon assimilation by eurotialean and hypocrealean fungi: Roles for CYP52 and CYP53 clans of cytochrome P450 genes," *Appl. Biochem. Biotechnol.*, vol. 184, pp. 1047–1060, 2018.
[69] S. Zibek, W. Wagner, T. Hirth, S. Rupp, and S. Huf, "Fermentative herstellung der α,ω-dicarbonsäure 1,18-oktadecendisäure als grundbaustein für biobasierte kunststoffe,"

Chemie Ing. Tech., vol. 81, no. 11, pp. 1797–1808, 2009.

[70] J.-M. Nicaud, F. Thevenieau, M.-T. Le Dall, and R. Marchal, "Production of dicarboxylic acids by improved mutant strains of *Yarrowia lipolytica*," US 2010/0041115 A1, 2010.

[71] L. Hammerer, C. K. Winkler, and W. Kroutil, "Regioselective biocatalytic hydroxylation of fatty acids by cytochrome P450s," *Catal. Letters*, vol. 148, pp. 787–812, 2018.

[72] P. Durairaj *et al.*, "Fungal cytochrome P450 monooxygenases of *Fusarium oxysporum* for the synthesis of ω-hydroxy fatty acids in engineered *Saccharomyces cerevisiae*," *Microb. Cell Fact.*, vol. 14, no. 45, pp. 1–16, 2015.

[73] M. Novak, L. Lah, M. Šala, J. Stojan, J. Bohlmann, and R. Komel, "Oleic acid metabolism via a conserved cytochrome P450 system-mediated ω- hydroxylation in the bark beetle-associated fungus *Grosmannia clavigera*," *PLoS One*, vol. 10, no. 3, p. e0120119, 2015.

[74] M. H. Hsu, Ü. Savas, K. J. Griffin, and E. F. Johnson, "Human cytochrome P450 family 4 enzymes: Function, genetic variation and regulation," *Drug Metab. Rev.*, vol. 39, pp. 515–538, 2007.

[75] F. Wernig, E. Boles, and M. Oreb, "De novo biosynthesis of 8-hydroxyoctanoic acid via a medium-chain length specific fatty acid synthase and cytochrome P450 in *Saccharomyces cerevisiae*," *Metab. Eng. Commun.*, vol. 10, no. e00111, 2020.

[76] F. Pinot and F. Beisson, "Cytochrome P450 metabolizing fatty acids in plants: Characterization and physiological roles," *FEBS J.*, vol. 278, pp. 195–205, 2011.

[77] R. Le Bouquin *et al.*, "CYP94A5, a new cytochrome P450 from *Nicotiana tabacum* is able to catalyze the oxidation of fatty acids to the ω-alcohol and to the corresponding diacid," *Eur. J. Biochem.*, vol. 268, pp. 3083–3090, 2001.

[78] S. Kandel *et al.*, "Characterization of a methyl jasmonate and wounding-responsive cytochrome P450 of *Arabidopsis thaliana* catalyzing dicarboxylic fatty acid formation in vitro," *FEBS J.*, vol. 274, pp. 5116–5127, 2007.

[79] L. Han, Y. Peng, Y. Zhang, W. Chen, Y. Lin, and Q. Wang, "Designing and creating a synthetic omega oxidation pathway in *Saccharomyces cerevisiae* enables production of medium-chain α, ω-dicarboxylic acids," *Front. Microbiol.*, vol. 8, no. 2184, pp. 1–12, 2017.

[80] T. Maier, H. H. Förster, O. Asperger, and U. Hahn, "Molecular characterization of the 56-kDa CYP153 from *Acinetobacter* sp. EB104," *Biochem. Biophys. Res. Commun.*, vol. 286, no. 3, pp. 652–658, 2001.

[81] J. B. Van Beilen *et al.*, "Cytochrome P450 alkane hydroxylases of the CYP153 family are common in alkane-degrading eubacteria lacking integral membrane alkane hydroxylases," *Appl. Environ. Microbiol.*, vol. 72, no. 1, pp. 59–65, 2006.

[82] S. H. Malca *et al.*, "Bacterial CYP153A monooxygenases for the synthesis of omega-hydroxylated fatty acids," *Chem. Commun.*, vol. 48, pp. 5115–5117, 2012.

[83] E. Jung *et al.*, "Production of ω-hydroxy palmitic acid using CYP153A35 and comparison of cytochrome P450 electron transfer system in vivo," *Appl. Microbiol. Biotechnol.*, vol. 100, pp. 10375–10384, 2016.

[84] J. B. Van Beilen, R. Holtackers, D. Lüscher, U. Bauer, B. Witholt, and W. A. Duetz, "Biocatalytic production of perillyl alcohol from limonene by using a novel *Mycobacterium* sp. cytochrome P450 alkane hydroxylase expressed in *Pseudomonas putida*," *Appl. Environ. Microbiol.*, vol. 71, no. 4, pp. 1737–1744, 2005.

[85] S. Cornelissen, M. K. Julsing, J. Volmer, O. Riechert, A. Schmid, and B. Bühler, "Whole-cell-based CYP153A6-catalyzed (S)-limonene hydroxylation efficiency depends on host background and profits from monoterpene uptake via AlkL," *Biotechnol. Bioeng.*, vol. 110, no. 5, pp. 1282–1292, 2013.

[86] M. Kubota *et al.*, "Isolation and functional analysis of cytochrome P450 CYP153A genes from various environments," *Biosci. Biotechnol. Biochem.*, vol. 69, no. 12, pp. 2421–2430, 2005.
[87] M. Bordeaux, A. Galarneau, F. Fajula, and J. Drone, "A regioselective biocatalyst for alkane activation under mild conditions," *Angew. Chemie - Int. Ed.*, vol. 50, pp. 2075–2079, 2011.
[88] D. Scheps, S. Honda Malca, S. M. Richter, K. Marisch, B. M. Nestl, and B. Hauer, "Synthesis of ω-hydroxy dodecanoic acid based on an engineered CYP153A fusion construct," *Microb. Biotechnol.*, vol. 6, no. 6, pp. 694–707, 2013.
[89] S. Notonier, Ł. Gricman, J. Pleiss, and B. Hauer, "Semirational protein engineering of CYP153AM.aq.-CPRBM3 for efficient terminal hydroxylation of short- to long-chain fatty acids," *ChemBioChem*, vol. 17, pp. 1550–1557, 2016.
[90] Y. Duan *et al.*, "Semi-rational engineering of cytochrome CYP153A from *Marinobacter aquaeolei* for improved ω-hydroxylation activity towards oleic acid," *Appl. Microbiol. Biotechnol.*, vol. 100, pp. 8779–8788, 2016.
[91] C. Sathesh-Prabu and S. K. Lee, "Production of long-chain α,ω-dicarboxylic acids by engineered *Escherichia coli* from renewable fatty acids and plant oils," *J. Agric. Food Chem.*, vol. 63, pp. 8199–8208, 2015.
[92] C. Sathesh-Prabu and S. K. Lee, "Enhancement of α,ω-dicarboxylic acid production by the expression of xylose reductase for refactoring redox cofactor regeneration," *J. Agric. Food Chem.*, vol. 66, pp. 3489–3497, 2018.
[93] C. H. Bowen, J. Bonin, A. Kogler, C. Barba-Ostria, and F. Zhang, "Engineering *Escherichia coli* for conversion of glucose to medium-chain ω-hydroxy fatty acids and α,ω-dicarboxylic acids," *ACS Synth. Biol.*, vol. 5, pp. 200–206, 2016.
[94] J. H. Bae, B. G. Park, E. Jung, P. G. Lee, and B. G. Kim, "fadD deletion and fadL overexpression in *Escherichia coli* increase hydroxy long-chain fatty acid productivity," *Appl. Microbiol. Biotechnol.*, vol. 98, pp. 8917–8925, 2014.
[95] E. Jung, B. G. Park, H. W. Yoo, J. Kim, K. Y. Choi, and B. G. Kim, "Semi-rational engineering of CYP153A35 to enhance ω-hydroxylation activity toward palmitic acid," *Appl. Microbiol. Biotechnol.*, vol. 102, pp. 269–277, 2018.
[96] H. A. Park and K. Y. Choi, "α, ω-Oxyfunctionalization of C12 alkanes via whole-cell biocatalysis of CYP153A from *Marinobacter aquaeolei* and a new CYP from *Nocardia farcinica* IFM10152," *Biochem. Eng. J.*, vol. 156, no. 107524, pp. 1–8, 2020.
[97] L. O. Narhi and A. J. Fulco, "Characterization of a catalytically self-sufficient 119,00-dalton cytochrome P-450 monooxygenase induced by barbiturates in *Bacillus megaterium*," *J. Biol. Chem.*, vol. 261, no. 16, pp. 7160–7169, 1986.
[98] C. J. C. Whitehouse, S. G. Bell, and L. L. Wong, "P450 BM3 (CYP102A1): Connecting the dots," *Chem. Soc. Rev.*, vol. 41, pp. 1218–1260, 2012.
[99] P. Meinhold, M. W. Peters, A. Hartwick, A. R. Hernandez, and F. H. Arnold, "Engineering cytochrome P450 BM3 for terminal alkane hydroxylation," *Adv. Synth. Catal.*, vol. 348, pp. 763–772, 2006.
[100] F. Brühlmann *et al.*, "Engineering cytochrome P450 BM3 of *Bacillus megaterium* for terminal oxidation of palmitic acid," *J. Biotechnol.*, vol. 184, pp. 17–26, 2014.
[101] J. B. Van Beilen and E. G. Funhoff, "Alkane hydroxylases involved in microbial alkane degradation," *Appl. Microbiol. Biotechnol.*, vol. 74, pp. 13–21, 2007.
[102] M. Kadisch *et al.*, "Maximization of cell viability rather than biocatalyst activity improves whole-cell ω-oxyfunctionalization performance," *Biotechnol. Bioeng.*, vol. 114, no. 4, pp. 874–884, 2017.
[103] D. Garfinkel, "Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions," *Arch. Biochem. Biophys.*, vol. 77, pp. 493–509, 1958.

[104] M. Klingenberg, “Pigments of rat liver microsomes,” *Arch. Biochem. Biophys.*, vol. 75, pp. 376–386, 1958.
[105] T. Omura and R. Sato, “The carbon monoxide-binding pigment of liver microsomes,” *J. Biol. Chem.*, vol. 239, no. 7, pp. 2370–2378, 1964.
[106] T. L. Poulos, B. C. Finzel, and A. J. Howard, “High-resolution crystal structure of cytochrome P450cam,” *J. Mol. Biol.*, vol. 195, pp. 687–700, 1987.
[107] F. P. Guengerich, “Human cytochrome P450 enzymes,” in *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 4th ed., P. R. Ortiz De Montellano, Ed. Springer International Publishing, 2015, pp. 523–785.
[108] J. A. Williams *et al.*, “Drug-drug interactions for UDP-glucuronosyltransferase substrates: A pharmacokinetic explanation for typically observed low exposure (AUC 1/AUC) ratios,” *Drug Metab. Dispos.*, vol. 32, no. 11, pp. 1201–1208, 2004.
[109] L. C. Wienkers and T. G. Heath, “Predicting in vivo drug interactions from in vitro drug discovery data,” *Nat. Rev. Drug Discov.*, vol. 4, pp. 825–833, 2005.
[110] F. P. Guengerich, “Cytochrome P450 research and The Journal of Biological Chemistry,” *J. Biol. Chem.*, vol. 294, no. 5, pp. 1671–1680, 2019.
[111] D. R. Nelson, “Cytochrome P450 diversity in the tree of life,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1866, pp. 141–154, 2018.
[112] M. A. Schuler and D. Werck-Reichhart, “Functional Genomics of P450s,” *Annu. Rev. Plant Biol.*, vol. 54, pp. 629–667, 2003.
[113] M. A. Schuler, “P450s in plants, insects, and their fungal pathogens,” in *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 4th ed., P. R. Ortiz De Montellano, Ed. Springer International Publishing, 2015, pp. 409–449.
[114] P. Durairaj, J.-S. Hur, and H. Yun, “Versatile biocatalysis of fungal cytochrome P450 monooxygenases,” *Microb. Cell Fact.*, vol. 15, no. 125, pp. 1–16, 2016.
[115] K. J. McLean, D. Leys, and A. W. Munro, “Microbial cytochromes P450,” in *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 4th ed., P. R. Ortiz De Montellano, Ed. Springer International Publishing, 2015, pp. 261–407.
[116] A. Greule, J. E. Stok, J. J. De Voss, and M. J. Cryle, “Unrivalled diversity: the many roles and reactions of bacterial cytochromes P450 in secondary metabolism,” *Nat. Prod. Rep.*, vol. 35, pp. 757–791, 2018.
[117] V. B. Urlacher and M. Girhard, “Cytochrome P450 monooxygenases in biotechnology and synthetic biology,” *Trends Biotechnol.*, vol. 37, no. 8, pp. 882–897, 2019.
[118] F. P. Guengerich and A. W. Munro, “Unusual cytochrome P450 enzymes and reactions,” *J. Biol. Chem.*, vol. 288, no. 24, pp. 17065–17073, 2013.
[119] J. A. McIntosh, C. C. Farwell, and F. H. Arnold, “Expanding P450 catalytic reaction space through evolution and engineering,” *Curr. Opin. Chem. Biol.*, vol. 19, pp. 126–134, 2014.
[120] M. T. Lundemo and J. M. Woodley, “Guidelines for development and implementation of biocatalytic P450 processes,” *Appl. Microbiol. Biotechnol.*, vol. 99, pp. 2465–2483, 2015.
[121] V. B. Urlacher and M. Girhard, “Cytochrome P450 monooxygenases: An update on perspectives for synthetic application,” *Trends Biotechnol.*, vol. 30, no. 1, pp. 26–36, 2012.
[122] R. Bernhardt and V. B. Urlacher, “Cytochromes P450 as promising catalysts for biotechnological application: Chances and limitations,” *Appl. Microbiol. Biotechnol.*, vol. 98, pp. 6185–6203, 2014.
[123] H. M. Girvan and A. W. Munro, “Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology,” *Curr. Opin. Chem. Biol.*, vol. 31, pp. 136–145, 2016.

[124] S. Notonier, M. Alexander, and L. N. Jayakody, “An overview of P450 enzymes: opportunity and challenges in industrial applications,” *Enzym. Eng.*, vol. 5, no. 3, pp. 1–5, 2016.
[125] K. Schroer, M. Kittelmann, and S. Lütz, “Recombinant human cytochrome P450 monooxygenases for drug metabolite synthesis,” *Biotechnol. Bioeng.*, vol. 106, no. 5, pp. 699–706, 2010.
[126] C. Duport, R. Spagnoli, E. Degryse, and D. Pompon, “Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast,” *Nat. Biotechnol.*, vol. 16, pp. 186–189, 1998.
[127] F. M. Szczebara *et al.*, “Total biosynthesis of hydrocortisone from a simple carbon source in yeast,” *Nat. Biotechnol.*, vol. 21, pp. 143–149, 2003.
[128] P. K. Ajikumar *et al.*, “Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*,” *Science.*, vol. 330, pp. 70–74, 2010.
[129] B. W. Biggs *et al.*, “Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in *Escherichia coli*,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 12, pp. 3209–3214, 2016.
[130] T. Matsuoka, S. Miyakoshi, K. Tanzawa, K. Nakahara, M. Hosobuchi, and N. Serizawa, “Purification and characterization of cytochrome P-450sca from *Streptomyces carbophilus*,” *Eur. J. Biochem.*, vol. 184, pp. 707–713, 1989.
[131] K. J. McLean *et al.*, “Single-step fermentative production of the cholesterol-lowering drug pravastatin via reprogramming of *Penicillium chrysogenum*,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 9, pp. 2847–2852, 2015.
[132] C. J. Paddon *et al.*, “High-level semi-synthetic production of the potent antimalarial artemisinin,” *Nature*, vol. 496, pp. 528–532, 2013.
[133] G. P. Moss, “Enzyme Nomenclature,” 2020. [Online]. Available: https://www.qmul.ac.uk/sbcs/iubmb/enzyme/. [Accessed: 09-Apr-2020].
[134] F. Hannemann, A. Bichet, K. M. Ewen, and R. Bernhardt, “Cytochrome P450 systems-biological variations of electron transport chains,” *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1770, pp. 330–344, 2007.
[135] D. R. Nelson, “The cytochrome P450 homepage,” *Hum. Genomics*, vol. 4, no. 1, pp. 59–65, 2009.
[136] L. Waskell and J. J. P. Kim, “Electron transfer partners of cytochrome P450,” in *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 4th ed., P. R. Ortiz De Montellano, Ed. Springer International Publishing, 2015, pp. 33–68.
[137] A. Ciaramella, D. Minerdi, and G. Gilardi, “Catalytically self-sufficient cytochromes P450 for green production of fine chemicals,” *Rend. Lincei*, vol. 28, no. Suppl 1, pp. S169–S181, 2017.
[138] S. Li, L. Du, and R. Bernhardt, “Redox partners: function modulators of bacterial P450 enzymes,” *Trends Microbiol.*, vol. 28, no. 6, pp. 445–454, 2020.
[139] T. D. Porter, “The roles of cytochrome *b5* in cytochrome P450 reactions,” *J. Biochem. Mol. Toxicol.*, vol. 16, no. 6, pp. 311–316, 2002.
[140] R. L. Wright, K. Harris, B. Solow, R. H. White, and P. J. Kennelly, “Cloning of a potential cytochrome p450 from the archaeon *Sulfolobus solfataricus*,” *FEBS Lett.*, vol. 384, pp. 235–239, 1996.
[141] K. L. Harris, R. E. S. Thomson, S. J. Strohmaier, Y. Gumulya, and E. M. J. Gillam, “Determinants of thermostability in the cytochrome P450 fold,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1866, pp. 97–115, 2018.
[142] C. J. Jackson *et al.*, “A novel sterol 14α-demethylase/ferredoxin fusion protein (MCCYP51FX) from *Methylococcus capsulatus* represents a new class of the cytochrome P450 superfamily,” *J. Biol. Chem.*, vol. 277, no. 49, pp. 46959–46965, 2002.

[143] C. S. Chong *et al.*, “Analysis of the xplAB-containing gene cluster involved in the bacterial degradation of the explosive hexahydro-1,3,5-trinitro-1,3,5- triazine,” *Appl. Environ. Microbiol.*, vol. 80, no. 21, pp. 6601–6610, 2014.
[144] G. A. Roberts, G. Grogan, A. Greter, S. L. Flitsch, and N. J. Turner, “Identification of a new class of cytochrome P450 from a *Rhodococcus* sp.,” *J. Bacteriol.*, vol. 184, no. 14, pp. 3898–3908, 2002.
[145] N. Nakayama, A. Takemae, and H. Shoun, “Cytochrome P450foxy, a catalytically self-sufficient fatty acid hydroxylase of the fungus *Fusarium oxysporum*,” *J. Biochem.*, vol. 119, no. 3, pp. 435–440, 1996.
[146] T. L. Poulos and E. F. Johnson, “Structures of cytochrome P450 enzymes,” in *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 4th ed., P. R. Ortiz De Montellano, Ed. Springer International Publishing, 2015, pp. 3–32.
[147] I. Schlichting *et al.*, “The catalytic pathway of cytochrome P450cam at atomic resolution,” *Science.*, vol. 287, pp. 1615–1622, 2000.
[148] S. Nagano and T. L. Poulos, “Crystallographic study on the dioxygen complex of wild-type and mutant cytochrome P450cam: Implications for the dioxygen activation mechanism,” *J. Biol. Chem.*, vol. 280, no. 36, pp. 31659–31663, 2005.
[149] T. L. Poulos, B. C. Finzel, and A. J. Howard, “Crystal structure of substrate-free *Pseudomonas putida* cytochrome P-450,” *Biochemistry*, vol. 25, no. 18, pp. 5314–5322, 1986.
[150] Y. T. Lee, R. F. Wilson, I. Rupniewski, and D. B. Goodin, “P450cam visits an open conformation in the absence of substrate,” *Biochemistry*, vol. 49, no. 16, pp. 3412–3419, 2010.
[151] O. Gotoh, “Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences,” *J. Biol. Chem.*, vol. 267, no. 1, pp. 83–90, 1992.
[152] J. A. Peterson and S. E. Graham, “A close family resemblance: The importance of structure in understanding cytochromes P450,” *Structure*, vol. 6, no. 9, pp. 1079–1085, 1998.
[153] K. G. Ravichandran, S. S. Boddupalli, C. A. Hasemann, J. A. Peterson, and J. Deisenhofer, “Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450’s,” *Science.*, vol. 261, pp. 731–736, 1993.
[154] H. Li and T. L. Poulos, “The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid,” *Nat. Struct. Biol.*, vol. 4, no. 2, pp. 140–146, 1997.
[155] D. Sehnal, A. S. Rose, J. Koča, S. K. Burley, and S. Velankar, “Mol*: Towards a common library and tools for web molecular graphics,” *Work. Mol. Graph. Vis. Anal. Mol. Data*, 2018.
[156] C. A. Hasemann, K. G. Ravichandran, J. A. Peterson, and J. Deisenhofer, “Crystal structure and refinement of cytochrome P450terp at 2·3 Å resolution,” *J. Mol. Biol.*, vol. 236, no. 4, pp. 1169–1185, 1994.
[157] J. R. Cupp-Vickery, H. Li, and T. L. Poulos, “Preliminary crystallographic analysis of an enzyme involved in erythromycin biosynthesis: Cytochrome P450eryF,” *Proteins Struct. Funct. Genet.*, vol. 20, no. 2, pp. 197–201, 1994.
[158] S. Y. Park *et al.*, “Crystal structure of nitric oxide reductase from denitrifying fungus *Fusarium oxysporum*,” *Nat. Struct. Biol.*, vol. 4, no. 10, pp. 827–832, 1997.
[159] P. A. Williams, J. Cosme, V. Sridhar, E. F. Johnson, and D. E. McRee, “Mammalian microsomal cytochrome P450 monooxygenase: Structural adaptations for membrane binding and functional diversity,” *Mol. Cell*, vol. 5, no. 1, pp. 121–131, 2000.
[160] E. F. Johnson and C. D. Stout, “Structural diversity of eukaryotic membrane cytochrome

P450s," *J. Biol. Chem.*, vol. 288, pp. 17082–17092, 2013.
[161] C. De Lemos-Chiarandini, A. B. Frey, D. D. Sabatini, and G. Kreibich, "Determination of the membrane topology of the phenobarbital-inducible rat liver cytochrome P-450 isoenzyme PB-4 using site-specific antibodies.," *J. Cell Biol.*, vol. 104, pp. 209–219, 1987.
[162] S. D. Black, S. T. Martin, and C. A. Smith, "Membrane topology of liver microsomal cytochrome P450 2B4 determined via monoclonal antibodies directed to the halt-transfer signal," *Biochemistry*, vol. 33, no. 22, pp. 6945–6951, 1994.
[163] M. L. Shank-Retzlaff, G. M. Raner, M. J. Coon, and S. G. Sligar, "Membrane topology of cytochrome P450 2B4 in Langmuir-Blodgett monolayers," *Arch. Biochem. Biophys.*, vol. 359, no. 1, pp. 82–88, 1998.
[164] T. H. Bayburt and S. G. Sligar, "Single-molecule height measurements on microsomal cytochrome P450 in nanometer-scale phospholipid bilayer disks," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 10, pp. 6725–6730, 2002.
[165] B. C. Monk *et al.*, "Architecture of a single membrane spanning cytochrome P450 suggests constraints that orient the catalytic domain relative to a bilayer," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 10, pp. 3865–3870, 2014.
[166] P. Urban, T. Lautier, D. Pompon, and G. Truan, "Ligand access channels in cytochrome P450 enzymes: A review," *Int. J. Mol. Sci.*, vol. 19, no. 1617, pp. 1–21, 2018.
[167] C. Barnaba, K. Gentry, N. Sumangala, and A. Ramamoorthy, "The catalytic function of cytochrome P450 is entwined with its membrane-bound nature," *F1000Research*, vol. 6, no. 662, pp. 1–10, 2017.
[168] M. Šrejber *et al.*, "Membrane-attached mammalian cytochromes P450: An overview of the membrane's effects on structure, drug binding, and interactions with redox partners," *J. Inorg. Biochem.*, vol. 183, pp. 117–136, 2018.
[169] F. P. Guengerich, M. R. Waterman, and M. Egli, "Recent structural insights into cytochrome P450 function," *Trends Pharmacol. Sci.*, vol. 37, no. 8, pp. 625–640, 2016.
[170] S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar, and W. Thiel, "P450 enzymes: Their structure, reactivity, and selectivity - modeled by QM/MM calculations," *Chem. Rev.*, vol. 110, no. 2, pp. 949–1017, 2010.
[171] K. D. Dubey and S. Shaik, "Cytochrome P450 - The wonderful nanomachine revealed through dynamic simulations of the catalytic cycle," *Acc. Chem. Res.*, vol. 52, pp. 389–399, 2019.
[172] I. G. Denisov, T. M. Makris, S. G. Sligar, and I. Schlichting, "Structure and chemistry of cytochrome P450," *Chem. Rev.*, vol. 105, no. 6, pp. 2253–2277, 2005.
[173] B. Meunier, S. P. de Visser, and S. Shaik, "Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes," *Chem. Rev.*, vol. 104, no. 9, pp. 3947–3980, 2004.
[174] J. T. Groves and G. A. McClusky, "Aliphatic hydroxylation via oxygen rebound. Oxygen transfer catalyzed by iron," *J. Am. Chem. Soc.*, vol. 98, no. 3, pp. 859–861, 1976.
[175] X. Huang and J. T. Groves, "Beyond ferryl-mediated hydroxylation: 40 years of the rebound mechanism and C–H activation," *J. Biol. Inorg. Chem.*, vol. 22, pp. 185–207, 2017.
[176] M. E. Albertolle and P. F. Guengerich, "The relationships between cytochromes P450 and H_2O_2: Production, reaction, and inhibition," *J. Inorg. Biochem.*, vol. 186, pp. 228–234, 2018.
[177] P. J. Mak and I. G. Denisov, "Spectroscopic studies of the cytochrome P450 reaction mechanisms," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1866, no. 1, pp. 178–204, 2018.
[178] R. Davydov, R. Kappl, J. Hüttermann, and J. A. Peterson, "EPR-spectroscopy of reduced oxyferrous-P450cam," *FEBS Lett.*, vol. 295, no. 1–3, pp. 113–115, 1991.

[179] J. Rittle and M. T. Green, “Cytochrome P450 compound I: Capture, characterization, and C-H bond activation kinetics,” *Science.*, vol. 330, pp. 933–937, 2010.
[180] F. P. Guengerich, “Mechanisms of cytochrome P450-catalyzed oxidations,” *ACS Catal.*, vol. 8, pp. 10964–10976, 2018.
[181] T. D. Porter and C. B. Kasper, “NADPH-cytochrome p-450 oxidoreductase: Flavin mononucleotide and flavin adenine dinucleotide domains evolved from different flavoproteins,” *Biochemistry*, vol. 25, no. 7, pp. 1682–1687, 1986.
[182] J. L. Vermilion and M. J. Coon, “Identification of the high and low potential flavins of liver microsomal NADPH-cytochrome P-450 reductase,” *J. Biol. Chem.*, vol. 253, no. 24, pp. 8812–8819, 1978.
[183] A. W. Munro, M. A. Noble, L. Robledo, S. N. Daff, and S. K. Chapman, “Determination of the redox properties of human NADPH-cytochrome P450 reductase,” *Biochemistry*, vol. 40, no. 7, pp. 1956–1963, 2001.
[184] T. Iyanagi and H. S. Mason, “Some properties of hepatic reduced nicotinamide adenine dinucleotide phosphate-cytochrome c reductase,” *Biochemistry*, vol. 12, no. 12, pp. 2297–2308, 1973.
[185] T. Iyanagi, R. Makino, and F. K. Anan, “Studies on the microsomal mixed-function oxidase system: Mechanism of action of hepatic NADPH-cytochrome P-450 reductase,” *Biochemistry*, vol. 20, no. 7, pp. 1722–1730, 1981.
[186] J. L. Vermilion and M. J. Coon, “Purified liver microsomal NADPH-cytochrome P-450 reductase. Spectral characterization of oxidation-reduction states,” *J. Biol. Chem.*, vol. 253, no. 8, pp. 2694–2704, 1978.
[187] M. B. Murataliev, R. Feyereisen, and F. A. Walker, “Electron transfer by diflavin reductases,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1698, pp. 1–26, 2004.
[188] I. Sevrioukova, C. Shaffer, D. P. Ballou, and J. A. Peterson, “Equilibrium and transient state spectrophotometric studies of the mechanism of reduction of the flavoprotein domain of P450BM-3,” *Biochemistry*, vol. 35, no. 22, pp. 7058–7068, 1996.
[189] M. B. Murataliev, M. Klein, A. Fulco, and R. Feyereisen, “Functional interactions in cytochrome P450BM3: Flavin semiquinone intermediates, role of NADP(H), and mechanism of electron transfer by the flavoprotein domain,” *Biochemistry*, vol. 36, no. 27, pp. 8401–8412, 1997.
[190] I. F. Sevrioukova, H. Li, H. Zhang, J. A. Peterson, and T. L. Poulos, “Structure of a cytochrome P450 – redox partner electron-transfer complex,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 96, pp. 1863–1868, 1999.
[191] M. Wang, D. L. Roberts, R. Paschke, T. M. Shea, B. S. S. Masters, and J.-J. P. Kim, “Three-dimensional structure of NADPH – cytochrome P450 reductase : Prototype for FMN- and FAD-containing enzymes,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 94, pp. 8411–8416, 1997.
[192] P. A. Hubbard, A. L. Shen, R. Paschke, C. B. Kasper, and J. J. P. Kim, “NADPH-cytochrome P450 oxidoreductase. Structural basis for hydride and electron transfer,” *J. Biol. Chem.*, vol. 276, no. 31, pp. 29163–29170, 2001.
[193] D. C. Lamb *et al.*, “A second FMN binding site in yeast NADPH-cytochrome P450 reductase suggests a mechanism of electron transfer by diflavin reductases,” *Structure*, vol. 14, pp. 51–61, 2006.
[194] C. Xia, S. P. Panda, C. C. Marohnic, P. Martaśek, B. S. Masters, and J. J. P. Kim, “Structural basis for human NADPH-cytochrome P450 oxidoreductase deficiency,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 33, pp. 13486–13491, 2011.
[195] A. C. Ebrecht, N. van der Bergh, S. T. L. Harrison, M. S. Smit, B. T. Sewell, and D. J. Opperman, “Biochemical and structural insights into the cytochrome P450 reductase from *Candida tropicalis*,” *Sci. Rep.*, vol. 9, no. 20088, pp. 1–11, 2019.

[196] P. Hlavica, J. Schulze, and D. F. V. Lewis, “Functional interaction of cytochrome P450 with its redox partners: A critical assessment and update of the topology of predicted contact regions,” *J. Inorg. Biochem.*, vol. 96, pp. 279–297, 2003.
[197] D. F. Estrada, J. S. Laurence, and E. E. Scott, “Substrate-modulated cytochrome P450 17A1 and cytochrome b5 interactions revealed by NMR,” *J. Biol. Chem.*, vol. 288, no. 23, pp. 17008–17018, 2013.
[198] H. L. Lin, C. Kenaan, H. Zhang, and P. F. Hollenberg, “Reaction of human cytochrome P450 3A4 with peroxynitrite: Nitrotyrosine formation on the proximal side impairs its interaction with NADPH-cytochrome P450 reductase,” *Chem. Res. Toxicol.*, vol. 25, pp. 2642–2653, 2012.
[199] L. Nikfarjam, S. Izumi, T. Yamazaki, and S. Kominami, “The interaction of cytochrome P450 17α with NADPH-cytochrome P450 reductase, investigated using chemical modification and MALDI-TOF mass spectrometry,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1764, pp. 1126–1131, 2006.
[200] C. Kenaan, H. Zhang, E. V. Shea, and P. F. Hollenberg, “Uncovering the role of hydrophobic residues in cytochrome P450-cytochrome P450 reductase interactions,” *Biochemistry*, vol. 50, pp. 3957–3967, 2011.
[201] H. Zhang, S. C. Im, and L. Waskell, “Cytochrome *b5* increases the rate of product formation by cytochrome P450 2B4 and competes with cytochrome P450 reductase for a binding site on cytochrome P450 2B4,” *J. Biol. Chem.*, vol. 282, no. 41, pp. 29766–29776, 2007.
[202] D. Hamdane, C. Xia, S. C. Im, H. Zhang, J. J. P. Kim, and L. Waskell, “Structure and function of an NADPH-cytochrome P450 oxidoreductase in an open conformation capable of reducing cytochrome P450,” *J. Biol. Chem.*, vol. 284, no. 17, pp. 11374–11384, 2009.
[203] A. Sündermann and C. Oostenbrink, “Molecular dynamics simulations give insight into the conformational change, complex formation, and electron transfer pathway for cytochrome P450 reductase,” *Protein Sci.*, vol. 22, pp. 1183–1195, 2013.
[204] D. F. Estrada, J. S. Laurence, and E. E. Scott, “Cytochrome P450 17A1 interactions with the FMN domain of its reductase as characterized by NMR,” *J. Biol. Chem.*, vol. 291, no. 8, pp. 3990–4003, 2016.
[205] F. Esteves *et al.*, “The Role of the FMN-domain of human cytochrome P450 oxidoreductase in its promiscuous interactions with structurally diverse redox partners,” *Front. Pharmacol.*, vol. 11, no. 299, pp. 1–16, 2020.
[206] L. Aigrain, D. Pompon, S. Moréra, and G. Truan, “Structure of the open conformation of a functional chimeric NADPH cytochrome P450 reductase,” *EMBO Rep.*, vol. 10, no. 7, pp. 742–747, 2009.
[207] J. Ellis, A. Gutierrez, I. L. Barsukov, W. C. Huang, J. G. Grossman, and G. C. K. Roberts, “Domain motion in cytochrome P450 reductase. Conformational equilibria revealed by NMR and small-angle x-ray scattering,” *J. Biol. Chem.*, vol. 284, no. 52, pp. 36628–36637, 2009.
[208] C. Xia *et al.*, “Conformational changes of NADPH-cytochrome P450 oxidoreductase are essential for catalysis and cofactor binding,” *J. Biol. Chem.*, vol. 286, no. 18, pp. 16246–16260, 2011.
[209] S. D. Black and M. J. Coon, “Structural features of liver microsomal NADPH-cytochrome P-450 reductase. Hydrophobic domain, hydrophilic domain, and connecting region.,” *J. Biol. Chem.*, vol. 257, no. 10, pp. 5929–5938, 1982.
[210] R. Huang *et al.*, “Probing the transmembrane structure and dynamics of microsomal nadph-cytochrome p450 oxidoreductase by solid-state NMR,” *Biophys. J.*, vol. 106, no. 10, pp. 2126–2133, 2014.

[211] A. Das and S. G. Sligar, “Modulation of the cytochrome P450 reductase redox potential by the phospholipid bilayer,” *Biochemistry*, vol. 48, no. 51, pp. 12104–12112, 2009.
[212] B. Louërat-Oriou, A. Perret, and D. Pompon, “Differential redox and electron-transfer properties of purified yeast, plant and human NADPH-cytochrome P-450 reductases highly modulate cytochrome P-450 activities,” *Eur. J. Biochem.*, vol. 258, pp. 1040–1049, 1998.
[213] D. A. Gideon, R. Kumari, A. M. Lynn, and K. M. Manoj, “What is the functional role of N-terminal transmembrane helices in the metabolism mediated by liver microsomal cytochrome P450 and its reductase?,” *Cell Biochem. Biophys.*, vol. 63, pp. 35–45, 2012.
[214] M. Miyamoto *et al.*, “Membrane anchor of cytochrome P450 reductase suppresses the uncoupling of cytochrome P450,” *Chem. Pharm. Bull.*, vol. 63, no. 4, pp. 286–294, 2015.
[215] C. Barnaba, M. J. Martinez, E. Taylor, A. O. Barden, and J. A. Brozik, “Single-protein tracking reveals that NADPH mediates the insertion of cytochrome P450 reductase into a biomimetic of the endoplasmic reticulum,” *J. Am. Chem. Soc.*, vol. 139, pp. 5420–5430, 2017.
[216] C. Xia *et al.*, “Structural and functional studies of the membrane-binding domain of NADPH-cytochrome P450 pxidoreductase,” *Biochemistry*, vol. 58, pp. 2408–2418, 2019.
[217] M. J. Gómez-Lechón, L. Tolosa, and M. T. Donato, “Upgrading HepG2 cells with adenoviral vectors that encode drug-metabolizing enzymes: application for drug hepatotoxicity testing,” *Expert Opin. Drug Metab. Toxicol.*, vol. 13, no. 2, pp. 137–148, 2017.
[218] M. Donato, A. Lahoz, J. Castell, and M. Gómez-Lechón, “Cell lines: A tool for in vitro drug metabolism studies,” *Curr. Drug Metab.*, vol. 9, no. 1, pp. 1–11, 2008.
[219] M. X. Zuber, E. R. Simpson, and M. R. Waterman, “Expression of bovine 17α-hydroxylase cytochrome P-450 cDNA in nonsteroidogenic (COS 1) cells,” *Science.*, vol. 234, pp. 1258–1261, 1986.
[220] Y. Katsumoto *et al.*, “Engineering of the rose flavonoid biosynthetic pathway successfully generated blue-hued flowers accumulating delphinidin,” *Plant Cell Physiol.*, vol. 48, no. 11, pp. 1589–1600, 2007.
[221] N. Shiota, A. Nagasawa, T. Sakaki, Y. Yabusaki, and H. Ohkawa, “Herbicide-resistant tobacco plants expressing the fused enzyme between rat cytochrome P4501A1 (CYP1A1) and yeast NADPH-cytochrome P450 oxidoreductase,” *Plant Physiol.*, vol. 106, pp. 17–23, 1994.
[222] S. Hirose, H. Kawahigashi, T. Inoue, H. Inui, H. Ohkawa, and Y. Ohkawa, “Enhanced expression of CYP2C9 and tolerance to sulfonylurea herbicides in transgenic rice plants,” *Plant Biotechnol.*, vol. 22, no. 2, pp. 89–96, 2005.
[223] K. Jensen, P. E. Jensen, and B. L. Møller, “Light-driven cytochrome P450 hydroxylations,” *ACS Chem. Biol.*, vol. 6, pp. 533–539, 2011.
[224] S. B. Mellor *et al.*, “Fusion of ferredoxin and cytochrome P450 enables direct light-driven biosynthesis,” *ACS Chem. Biol.*, vol. 11, pp. 1862–1869, 2016.
[225] A. Z. Nielsen *et al.*, “Redirecting photosynthetic reducing power toward bioactive natural product synthesis,” *ACS Synth. Biol.*, vol. 2, pp. 308–315, 2013.
[226] T. Gnanasekaran *et al.*, “Transfer of the cytochrome P450-dependent dhurrin pathway from *Sorghum bicolor* into *Nicotiana tabacum* chloroplasts for light-driven synthesis,” *J. Exp. Bot.*, vol. 67, no. 8, pp. 2495–2506, 2016.
[227] A. Wlodarczyk *et al.*, “Metabolic engineering of light-driven cytochrome P450 dependent pathways into *Synechocystis* sp. PCC 6803,” *Metab. Eng.*, vol. 33, pp. 1–11, 2016.
[228] H. Duan and M. A. Schuler, “Heterologous expression and strategies for encapsulation

of membrane-localized plant P450s," *Phytochem. Rev.*, vol. 5, pp. 507–523, 2006.
[229] J. Hausjell, H. Halbwirth, and O. Spadiut, "Recombinant production of eukaryotic cytochrome P450s in microbial cell factories," *Biosci. Rep.*, vol. 38, pp. 1–13, 2018.
[230] S. Zelasko, A. Palaria, and A. Das, "Optimizations to achieve high-level expression of cytochrome P450 proteins using *Escherichia coli* expression systems," *Protein Expr. Purif.*, vol. 92, pp. 77–87, 2013.
[231] S. Li, Y. Li, and C. D. Smolke, "Strategies for microbial synthesis of high-value phytochemicals," *Nat. Chem.*, vol. 10, pp. 395–404, 2018.
[232] K. Oeda, T. Sakaki, and H. Ohkawa, "Expression of rat liver cytochrome P-450MC cDNA in *Saccharomyces cerevisiae*," *DNA*, vol. 4, no. 3, pp. 203–210, 1985.
[233] P. Urban *et al.*, "Characterization of recombinant plant cinnamate 4-hydroxylase produced in yeast. Kinetic and spectral properties of the major plant P450 of the phenylpropanoid pathway," *Eur. J. Biochem.*, vol. 222, pp. 843–850, 1994.
[234] J. Hausjell, D. Schendl, J. Weissensteiner, C. Molitor, H. Halbwirth, and O. Spadiut, "Recombinant production of a hard-to-express membrane-bound cytochrome P450 in different yeasts—Comparison of physiology and productivity," *Yeast*, vol. 37, pp. 217–226, 2020.
[235] P. Durairaj *et al.*, "Functional expression and activity screening of all human cytochrome P450 enzymes in fission yeast," *FEBS Lett.*, vol. 593, pp. 1372–1380, 2019.
[236] H. Ichinose, "Molecular and functional diversity of fungal cytochrome P450s," *Biol. Pharm. Bull.*, vol. 35, no. 6, pp. 833–837, 2012.
[237] G. A. Schoch, R. Attias, M. Belghazi, P. M. Dansette, and D. Werck-Reichhart, "Engineering of a water-soluble plant cytochrome P450, CYP73A1, and NMR-based orientation of natural and alternate substrates in the active site," *Plant Physiol.*, vol. 133, pp. 1198–1208, 2003.
[238] S. Galanie, K. Thodey, I. J. Trenchard, M. F. Interrante, and C. D. Smolke, "Complete biosynthesis of opioids in yeast," *Science.*, vol. 349, no. 6252, pp. 1095–1100, 2015.
[239] I. N. Harnastai, A. A. Gilep, and S. A. Usanov, "The development of an efficient system for heterologous expression of cytochrome P450s in *Escherichia coli* using hemA gene co-expression," *Protein Expr. Purif.*, vol. 46, pp. 47–55, 2006.
[240] H. J. Barnes, M. P. Arlotto, and M. R. Waterman, "Expression and enzymatic activity of recombinant cytochrome P450 17α-hydroxylase in *Escherichia coli*," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, pp. 5597–5601, 1991.
[241] M. C. Y. Chang, R. A. Eachus, W. Trieu, D. K. Ro, and J. D. Keasling, "Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s," *Nat. Chem. Biol.*, vol. 3, no. 5, pp. 274–277, 2007.
[242] M. P. Pritchard *et al.*, "A general strategy for the expression of recombinant human cytochrome P450s in *Escherichia coli* using bacterial signal peptides: Expression of CYP3A4, CYP2A6, and CYP2E1," *Arch. Biochem. Biophys.*, vol. 345, no. 2, pp. 342–354, 1997.
[243] U. Christensen *et al.*, "De-bugging and maximizing plant cytochrome P450 production in *Escherichia coli* with C-terminal GFP fusions," *Appl. Microbiol. Biotechnol.*, vol. 101, pp. 4103–4113, 2017.
[244] D. Vazquez-Albacete, A. M. Cavaleiro, U. Christensen, S. Seppälä, B. L. Møller, and M. H. H. Nørholm, "An expression tag toolbox for microbial production of membrane bound plant cytochromes P450," *Biotechnol. Bioeng.*, vol. 114, no. 4, pp. 751–760, 2017.
[245] H. Ichinose and H. Wariishi, "High-level heterologous expression of fungal cytochrome P450s in *Escherichia coli*," *Biochem. Biophys. Res. Commun.*, vol. 438, pp. 289–294, 2013.
[246] H. Ichinose, M. Hatakeyama, and Y. Yamauchi, "Sequence modifications and

heterologous expression of eukaryotic cytochromes P450 in *Escherichia coli*," *J. Biosci. Bioeng.*, vol. 120, no. 3, pp. 268–274, 2015.
[247] S. D. Schumacher and J. Jose, "Expression of active human P450 3A4 on the cell surface of *Escherichia coli* by Autodisplay," *J. Biotechnol.*, vol. 161, pp. 113–120, 2012.
[248] J. Jose and T. F. Meyer, "The Autodisplay story, from discovery to biotechnical and biomedical applications," *Microbiol. Mol. Biol. Rev.*, vol. 71, no. 4, pp. 600–619, 2007.
[249] P. Quehl, J. Hollender, J. Schüürmann, T. Brossette, R. Maas, and J. Jose, "Co-expression of active human cytochrome P450 1A2 and cytochrome P450 reductase on the cell surface of *Escherichia coli*," *Microb. Cell Fact.*, vol. 15, no. 26, pp. 1–15, 2016.
[250] S.-K. Yim *et al.*, "Surface display of heme-and diflavin-containing cytochrome P450 BM3 in *Escherichia coli*: A whole-cell biocatalyst for oxidation," *J. Microbiol. Biotechnol*, vol. 20, no. 4, pp. 712–717, 2010.
[251] P. K. Wolber, "Bacterial ice nucleation," *Adv. Microb. Physiol.*, vol. 34, pp. 203–237, 1993.
[252] E. Van Bloois, R. T. Winter, H. Kolmar, and M. W. Fraaije, "Decorating microbes: surface display of proteins on *Escherichia coli*," *Trends Biotechnol.*, vol. 29, no. 2, pp. 79–86, 2010.
[253] M. L. Pham and M. Polakovič, "Microbial cell surface display of oxidoreductases: Concepts and applications," *Int. J. Biol. Macromol.*, vol. 165, pp. 835–841, 2020.
[254] F. W. Ströhle, E. Kranen, J. Schrader, R. Maas, and D. Holtmann, "A simplified process design for P450 driven hydroxylation based on surface displayed enzymes," *Biotechnol. Bioeng.*, vol. 113, no. 6, pp. 1225–1233, 2016.
[255] S. Gallus, T. Peschke, M. Paulsen, T. Burgahn, C. M. Niemeyer, and K. S. Rabe, "Surface display of complex enzymes by in situ SpyCatcher-SpyTag interaction," *ChemBioChem*, vol. 21, no. 15, pp. 2126–2131, 2020.
[256] E. M. J. Gillam, "Engineering cytochrome P450 enzymes," *Chem. Res. Toxicol.*, vol. 21, no. 1, pp. 220–231, 2008.
[257] D. R. McDougle *et al.*, "Incorporation of charged residues in the CYP2J2 F-G loop disrupts CYP2J2-lipid bilayer interactions," *Biochim. Biophys. Acta - Biomembr.*, vol. 1848, pp. 2460–2470, 2015.
[258] C. Von Wachenfeldt, T. H. Richardson, J. Cosme, and E. F. Johnson, "Microsomal P450 2C3 is expressed as a soluble dimer in *Escherichia coli* following modifications of its N-terminus," *Arch. Biochem. Biophys.*, vol. 339, no. 1, pp. 107–114, 1997.
[259] J. Cosme and E. F. Johnson, "Engineering microsomal cytochrome P450 2C5 to be a soluble, monomeric enzyme," *J. Biol. Chem.*, vol. 275, no. 4, pp. 2545–2553, 2000.
[260] P. A. Williams, J. Cosme, A. Ward, H. C. Angove, D. M. Vinković, and H. Jhoti, "Crystal structure of human cytochrome P450 2C9 with bound warfarin," *Nature*, vol. 424, pp. 464–468, Jul. 2003.
[261] G. A. Schoch, J. K. Yano, M. R. Wester, K. J. Griffin, C. D. Stout, and E. F. Johnson, "Structure of human microsomal cytochrome P450 2C8: Evidence for a peripheral fatty acid binding site," *J. Biol. Chem.*, vol. 279, no. 10, pp. 9497–9503, 2004.
[262] J. K. Yano, M. H. Hsu, K. J. Griffin, C. D. Stout, and E. F. Johnson, "Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen," *Nat. Struct. Mol. Biol.*, vol. 12, no. 9, pp. 822–823, 2005.
[263] C. M. Jenkins and M. R. Waterman, "Flavodoxin and NADPH-flavodoxin reductase from *Escherichia coli* support bovine cytochrome P450c17 hydroxylase activities," *J. Biol. Chem.*, vol. 269, no. 44, pp. 27401–27408, 1994.
[264] I. Neunzig *et al.*, "Coexpression of CPR from various origins enhances biotransformation activity of human CYPs in *S. pombe*," *Appl. Biochem. Biotechnol.*, vol. 170, pp. 1751–1766, 2013.

[265] P. Durairaj, E. Jung, H. H. Park, B. G. Kim, and H. Yun, “Comparative functional characterization of a novel benzoate hydroxylase cytochrome P450 of *Fusarium oxysporum*,” *Enzyme Microb. Technol.*, vol. 70, pp. 58–65, 2015.
[266] T. S. Chang, S. Y. Chao, and Y. C. Chen, “Production of ortho-hydroxydaidzein derivatives by a recombinant strain of *Pichia pastoris* harboring a cytochrome P450 fusion gene,” *Process Biochem.*, vol. 48, pp. 426–429, 2013.
[267] W. Zhang *et al.*, “New reactions and products resulting from alternative interactions between the P450 enzyme and redox partners,” *J. Am. Chem. Soc.*, vol. 136, pp. 3640–3646, 2014.
[268] H. Murakami, Y. Yabusaki, T. Sakaki, M. Shibata, and H. Ohkawa, “A genetically engineerd P450 monooxygenase : construction of the functional fused enzyme between rat cytochrome P450c and NADPH-cytochrome P450 reductase,” *DNA*, vol. 6, no. 3, pp. 189–197, 1987.
[269] M. Shibata, T. Sakaki, Y. Yabusaki, H. Murakami, and H. Ohkawa, “Genetically engineered P450 monooxygenases: construction of bovine P450c17/yeast reductase fused enzymes,” *DNA Cell Biol.*, vol. 9, no. 1, pp. 27–36, 1990.
[270] C. W. Fisher, M. S. Shet, D. L. Caudle, C. A. Martin-Wixtrom, and R. W. Estabrook, “High-level expression in *Escherichia coli* of enzymatically active fusion proteins containing the domains of mammalian cytochromes P450 and NADPH-P450 reductase flavoprotein,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 89, pp. 10817–10821, 1992.
[271] S. J. Sadeghi and G. Gilardi, “Chimeric P450 enzymes: Activity of artificial redox fusions driven by different reductases for biotechnological applications,” *Biotechnol. Appl. Biochem.*, vol. 60, no. 1, pp. 102–110, 2013.
[272] H. Inui, A. Maeda, and H. Ohkawa, “Molecular characterization of specifically active recombinant fused enzymes consisting of CYP3A4, NADPH-cytochrome P450 oxidoreductase, and cytochrome *b5*,” *Biochemistry*, vol. 46, no. 35, pp. 10213–10221, 2007.
[273] G. Gilardi, Y. T. Meharenna, G. E. Tsotsou, S. J. Sadeghi, M. Fairhead, and S. Giannini, “Molecular Lego: Design of molecular assemblies of P450 enzymes for nanobiotechnology,” *Biosens. Bioelectron.*, vol. 17, pp. 133–145, 2002.
[274] D. Degregorio *et al.*, “Human cytochrome P450 3A4 as a biocatalyst : Effects of the engineered linker in modulation of coupling efficiency in 3A4-BMR chimeras,” *Front. Pharmacol.*, vol. 8, no. 121, pp. 1–13, 2017.
[275] S. Castrignanò, S. D’Avino, G. Di Nardo, G. Catucci, S. J. Sadeghi, and G. Gilardi, “Modulation of the interaction between human P450 3A4 and *B. megaterium* reductase via engineered loops,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1866, pp. 116–125, 2018.
[276] M. Nodate, M. Kubota, and N. Misawa, “Functional expression system for cytochrome P450 genes using the reductase domain of self-sufficient P450RhF from *Rhodococcus* sp. NCIMB 9784,” *Appl. Microbiol. Biotechnol.*, vol. 71, pp. 455–462, 2006.
[277] J. Schückel, E. L. Rylott, G. Grogan, and N. C. Bruce, “A gene-fusion approach to enabling plant cytochromes P450 for biocatalysis,” *ChemBioChem*, vol. 13, pp. 2758–2763, 2012.
[278] A. W. Munro, K. J. McLean, J. L. Grant, and T. M. Makris, “Structure and function of the cytochrome P450 peroxygenase enzymes,” *Biochem. Soc. Trans.*, vol. 46, pp. 183–196, 2018.
[279] M. R. Anari, P. D. Josephy, T. Henry, and P. J. O’Brien, “Hydrogen peroxide supports human and rat cytochrome P450 1A2-catalyzed 2-amino-3-methylimidazo[4,5-f]quinoline bioactivation to mutagenic metabolites: Significance of cytochrome P450 peroxygenase,” *Chem. Res. Toxicol.*, vol. 10, no. 5, pp. 582–588, 1997.

[280] A. Chefson, J. Zhao, and K. Auclair, "Replacement of natural cofactors by selected hydrogen peroxide donors or organic peroxides results in improved activity for CYP3A4 and CYP2D6," *ChemBioChem*, vol. 7, no. 6, pp. 916–919, 2006.
[281] Z. Zhang, Y. Li, R. A. Stearns, P. R. Ortiz de Montellano, T. A. Baillie, and W. Tang, "Cytochrome P450 3A4-mediated oxidative conversion of a cyano to an amide group in the metabolism of pinacidil," *Biochemistry*, vol. 41, no. 8, pp. 2712–2718, 2002.
[282] I. H. Hanna, J. A. Krauser, H. Cai, M. S. Kim, and F. P. Guengerich, "Diversity in mechanisms of substrate oxidation by cytochrome P450 2D6: Lack of an allosteric role of NADPH-cytochrome P450 reductase in catalytic regioselectmty," *J. Biol. Chem.*, vol. 276, no. 43, pp. 39553–39561, 2001.
[283] P. Cirino and F. Arnold, "A self-sufficient peroxide-driven hydroxylation biocatalyst," *Angew. Chemie*, vol. 115, pp. 3421–3423, 2003.
[284] P. Hlavica, "Assembly of non-natural electron transfer conduits in the cytochrome P450 system: A critical assessment and update of artificial redox constructs amenable to exploitation in biotechnological areas," *Biotechnol. Adv.*, vol. 27, pp. 103–121, 2009.
[285] S. J. Sadeghi, A. Fantuzzi, and G. Gilardi, "Breakthrough in P450 bioelectrochemistry and future perspectives," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1814, pp. 237–248, 2011.
[286] S. Krishnan, "Bioelectrodes for evaluating molecular therapeutic and toxicity properties," *Curr. Opin. Electrochem.*, vol. 19, pp. 20–26, 2020.
[287] H. Shalan, M. Kato, and L. Cheruzel, "Keeping the spotlight on cytochrome P450," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1866, pp. 80–87, 2018.
[288] H. Schewe, B. A. Kaup, and J. Schrader, "Improvement of P450BM-3 whole-cell biocatalysis by integrating heterologous cofactor regeneration combining glucose facilitator and dehydrogenase in E. coli," *Appl. Microbiol. Biotechnol.*, vol. 78, pp. 55–65, 2008.
[289] A. Siriphongphaew *et al.*, "Development of a whole-cell biocatalyst co-expressing P450 monooxygenase and glucose dehydrogenase for synthesis of epoxyhexane," *Appl. Microbiol. Biotechnol.*, vol. 95, pp. 357–367, 2012.
[290] V. B. Urlacher and S. Schulz, "Multi-enzyme systems and cascade reactions involving cytochrome P450 monooxygenases," in *Cascade Biocatalysis: Integrating Stereoselective and Environmentally Friendly Reactions*, S. Riva and W.-D. Fessner, Eds. Wiley-VCH Verlag GmbH & Co. KGaA, 2014, pp. 87–132.
[291] E. E. Ferrandi, D. Monti, and S. Riva, "New trends in the in situ enzymatic recycling of NAD(P)(H) cofactors," in *Cascade Biocatalysis: Integrating Stereoselective and Environmentally Friendly Reactions*, S. Riva and W.-D. Fessner, Eds. Wiley-VCH Verlag GmbH & Co. KGaA, 2014, pp. 23–42.
[292] T. Kubo, M. W. Peters, P. Meinhold, and F. H. Arnold, "Enantioselective epoxidation of terminal alkenes to (R)- and (S)-epoxides by engineered cytochromes P450 BM-3," *Chem. - A Eur. J.*, vol. 12, pp. 1216–1220, 2006.
[293] S. C. Maurer, H. Schulze, R. D. Schmid, and V. Urlacher, "Immobilisation of P450 BM-3 and an NADP+ cofactor recycling system: towards a technical application of heme-containing monooxygenases in fine chemical synthesis," *Adv. Synth. Catal.*, vol. 345, pp. 802–810, 2003.
[294] N. Beyer, J. K. Kulig, A. Bartsch, M. A. Hayes, D. B. Janssen, and M. W. Fraaije, "P450BM3 fused to phosphite dehydrogenase allows phosphite-driven selective oxidations," *Appl. Microbiol. Biotechnol.*, vol. 101, pp. 2319–2331, 2017.
[295] A. M. Chánique and L. P. Parra, "Protein engineering for nicotinamide coenzyme specificity in oxidoreductases: Attempts and challenges," *Front. Microbiol.*, vol. 9, no. 194, pp. 1–14, 2018.

[296] S. C. Maurer, K. Kühnel, L. A. Kaysser, S. Eiben, R. D. Schmid, and V. B. Urlacher, "Catalytic hydroxylation in biphasic systems using CYP102A1 mutants," *Adv. Synth. Catal.*, vol. 347, pp. 1090–1098, 2005.
[297] V. B. Urlacher, S. Lutz-Wahl, and R. D. Schmid, "Microbial P450 enzymes in biotechnology," *Appl. Microbiol. Biotechnol.*, vol. 64, pp. 317–325, 2004.
[298] D. Valikhani, J. M. Bolivar, A. Dennig, and B. Nidetzky, "A tailor-made, self-sufficient and recyclable monooxygenase catalyst based on coimmobilized cytochrome P450 BM3 and glucose dehydrogenase," *Biotechnol. Bioeng.*, vol. 115, pp. 2416–2425, 2018.
[299] S. B. Lamb, D. C. Lamb, S. L. Kelly, and D. C. Stuckey, "Cytochrome P450 immobilisation as a route to bioremediation/biocatalysis," *FEBS Lett.*, vol. 431, pp. 343–346, 1998.
[300] M. Ingelman-Sundberg and H. Glaumann, "Incorporation of purified components of the rabbit liver microsomal hydroxylase system into phospholipid vesicles," *Biochim. Biophys. Acta - Biomembr.*, vol. 599, pp. 417–435, 1980.
[301] T. Ahn, F. P. Guengerich, and C. H. Yun, "Membrane insertion of cytochrome P450 1A2 promoted by anionic phospholipids," *Biochemistry*, vol. 37, no. 37, pp. 12860–12866, 1998.
[302] A. Nath, W. M. Atkins, and S. G. Sligar, "Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins," *Biochemistry*, vol. 46, no. 8, pp. 2059–2069, 2007.
[303] I. G. Denisov and S. G. Sligar, "Cytochromes P450 in nanodiscs," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1814, pp. 223–229, 2011.
[304] I. G. Denisov and S. G. Sligar, "Nanodiscs in Membrane Biochemistry and Biophysics," *Chem. Rev.*, vol. 117, pp. 4669–4713, 2017.
[305] D. Jendrossek and R. Handrick, "Microbial degradation of polyhydroxyalkanoates," *Annu. Rev. Microbiol.*, vol. 56, pp. 403–432, 2002.
[306] J. H. Lee *et al.*, "New platform for cytochrome P450 reaction combining in situ immobilization on biopolymer," *Bioconjug. Chem.*, vol. 25, pp. 2101–2104, 2014.
[307] B. Stenger, A. Gerber, R. Bernhardt, and F. Hannemann, "Functionalized poly(3-hydroxybutyric acid) bodies as new in vitro biocatalysts," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1866, pp. 52–59, 2018.
[308] H. Ngo and T. A. Foglia, "Method of producing dicarboxylic acids," US 7,534,917 B1, 2009.
[309] R. A. Gross, W. Lu, J. Ness, and J. Minshull, "Production of an α-carboxyl-ω-hydroxy fatty acid using a genetically modified *Candida* strain," US 8,158,391 B2, 2012.
[310] F. He and Y. T. Chen, "Cloning and heterologous expression of the NADPH cytochrome P450 oxidoreductase genes from an industrial dicarboxylic acid-producing *Candida tropicalis*," *Yeast*, vol. 22, pp. 481–491, 2005.
[311] J. M. Cregg, J. L. Cereghino, J. Shi, and D. R. Higgins, "Recombinant protein expression in *Pichia pastoris*," *Appl. Biochem. Biotechnol. - Part B Mol. Biotechnol.*, vol. 16, no. 1, pp. 23–52, 2000.
[312] F. S. Hartner and A. Glieder, "Regulation of methanol utilisation pathway genes in yeasts," *Microb. Cell Fact.*, vol. 5, no. 39, pp. 1–21, 2006.
[313] H. Waegeman and W. Soetaert, "Increasing recombinant protein production in *Escherichia coli* through metabolic and genetic engineering," *J. Ind. Microbiol. Biotechnol.*, vol. 38, pp. 1891–1910, 2011.
[314] M. Ahmad, M. Hirz, H. Pichler, and H. Schwab, "Protein expression in *Pichia pastoris*: Recent achievements and perspectives for heterologous protein production," *Appl. Microbiol. Biotechnol.*, vol. 98, no. 12, pp. 5301–5317, 2014.
[315] M. Guerfal *et al.*, "The *HAC1* gene from *Pichia pastoris*: characterization and effect of

its overexpression on the production of secreted, surface displayed and membrane proteins," *Microb. Cell Fact.*, vol. 9, no. 49, pp. 1–12, 2010.
[316] K. Venkateswarlu, D. C. Lamb, D. E. Kelly, N. J. Manning, and S. L. Kelly, "The N-terminal membrane domain of yeast NADPH-cytochrome P450 (CYP) oxidoreductase is not required for catalytic activity in sterol biosynthesis or in reconstitution of CYP activity," *J. Biol. Chem.*, vol. 273, no. 8, pp. 4492–4496, 1998.
[317] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, and M. J. Sternberg, "The Phyre2 web portal for protein modeling, prediction and analysis," *Nat. Protoc.*, vol. 10, no. 6, pp. 845–858, 2016.
[318] J. Quan and J. Tian, "Circular polymerase extension cloning of complex gene libraries and pathways," *PLoS One*, vol. 4, no. 7, pp. 1–6, 2009.
[319] L. Zheng, U. Baumann, and J. L. Reymond, "An efficient one-step site-directed and site-saturation mutagenesis protocol.," *Nucleic Acids Res.*, vol. 32, no. 14, pp. 1–5, 2004.
[320] M. Lõoke, K. Kristjuahan, and A. Kristjuhan, "Extraction of genomic DNA From yeasts for PCR-based applications," *Biotechniques*, vol. 50, no. 5, pp. 325–328, 2011.
[321] D. Pompon, B. Louerat, A. Bronine, and P. Urban, "Yeast expression of animal and plant P450s in optimized redox environments," *Methods Enzymol.*, vol. 272, pp. 51–64, 1996.
[322] M. Gilewicz, M. Zacek, J.-C. Bertrand, and E. Azoulay, "Hydroxylase regulation in *Candida tropicalis* grown on alkanes," *Can. J. Microbiol.*, vol. 25, pp. 201–206, 1979.
[323] R. A. Gross, W. Lu, J. Ness, and J. Minshull, "Biosynthetic routes to long-chain alpha,omega-hydroxyacids, diacids and their conversion to oligomers and polymers," US 2010/0285545 A1, 2010.
[324] R. Gurdeep Singh *et al.*, "Unipept 4.0: Functional Analysis of Metaproteome Data," *J. Proteome Res.*, vol. 18, pp. 606–615, 2019.
[325] L. Klug *et al.*, "The lipidome and proteome of microsomes from the methylotrophic yeast *Pichia pastoris*," *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids*, vol. 1841, pp. 215–226, 2014.
[326] T. Wriessnegger *et al.*, "Lipid composition of peroxisomes from the yeast *Pichia pastoris* grown on different carbon sources," *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids*, vol. 1771, pp. 455–461, 2007.
[327] R. K. Gudiminchi, M. Geier, A. Glieder, and A. Camattari, "Screening for cytochrome P450 expression in *Pichia pastoris* whole cells by P450-carbon monoxide complex determination," *Biotechnol. J.*, vol. 8, pp. 146–152, 2013.
[328] C. Zhao, G. Song, H. Duan, T. Tang, C. Wang, and L. Qiu, "Heterologous expression of *Helicoverpa armigera* cytochrome P450 CYP6B7 in *Pichia pastoris* and interactions of CYP6B7 with insecticides," *Pest Manag. Sci.*, vol. 73, pp. 1866–1872, 2017.
[329] J. M. Trant, "Functional expression of recombinant spiny dogfish shark cytochrome P450c17 (17a-hydroxylase/C17,20-lyase) in yeast (*Pichia pastoris*)," *Arch. Biochem. Biophys.*, vol. 326, no. 1, pp. 8–14, 1996.
[330] N. W. Kolar, A. C. Swart, J. I. Mason, and P. Swart, "Functional expression and characterisation of human cytochrome P45017α in *Pichia pastoris*," *J. Biotechnol.*, vol. 129, pp. 635–644, 2007.
[331] P. Buathong, N. Boonvitthya, G. Truan, and W. Chulalaksananukul, "Biotransformation of lauric acid into 1,12-dodecanedioic acid using CYP52A17 expressed in *Saccharomyces cerevisiae* and its application in refining coconut factory wastewater," *Int. Biodeterior. Biodegrad.*, vol. 139, pp. 70–77, 2019.
[332] S. M. Boyle, M. P. Popp, W. C. Smith, R. M. Greenberg, and M. O. James, "Expression of CYP2L1 in the yeast *Pichia pastoris*, and determination of catalytic activity with progesterone and testosterone," *Mar. Environ. Res.*, vol. 46, no. 1–5, pp. 25–28, 1998.
[333] M. D. Andersen, P. K. Busk, I. Svendsen, and B. L. Møller, "Cytochromes P-450 from

cassava (Manihot esculenta Crantz) catalyzing the first steps in the biosynthesis of the cyanogenic glucosides linamarin and lotaustralin. Cloning, functional expression in *Pichia pastoris*, and substrate specificity of the isolated ," *J. Biol. Chem.*, vol. 275, no. 3, pp. 1966–1975, 2000.

[334] F. W. Krainer *et al.*, "Optimizing cofactor availability for the production of recombinant heme peroxidase in *Pichia pastoris*.," *Microb. Cell Fact.*, vol. 14, no. 4, pp. 1–9, 2015.

[335] M. Dietrich *et al.*, "Recombinant production of human microsomal cytochrome P450 2D6 in the methylotrophic yeast *Pichia pastoris*," *ChemBioChem*, vol. 6, pp. 2014–2022, 2005.

[336] M. Geier, C. Schmid, and A. Glieder, "First functional expression of cytochrome P450 3A4 in *Pichia pastoris*," *Chim. Oggi - Chemsitry Today*, vol. 31, no. 3, pp. 24–27, 2013.

[337] F. Matsuzaki and H. Wariishi, "Molecular characterization of cytochrome P450 catalyzing hydroxylation of benzoates from the white-rot fungus *Phanerochaete chrysosporium*," *Biochem. Biophys. Res. Commun.*, vol. 334, pp. 1184–1190, 2005.

[338] K. Syed, A. Porollo, Y. W. Lam, and J. S. Yadav, "A fungal P450 (CYP5136A3) capable of oxidizing polycyclic aromatic hydrocarbons and endocrine disrupting alkylphenols: Role of Trp129 and Leu324," *PLoS One*, vol. 6, no. 12, pp. 1–14, 2011.

[339] S. De Waele *et al.*, "Optimized expression of the *Starmerella bombicola* lactone esterase in *Pichia pastoris* through temperature adaptation, codon-optimization and co-expression with *HAC1*," *Protein Expr. Purif.*, vol. 143, pp. 62–70, 2018.

[340] F. P. Guengerich, M. V. Martin, C. D. Sohl, and Q. Cheng, "Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase," *Nat. Protoc.*, vol. 4, no. 9, pp. 1–7, 2009.

[341] K. Hayashi, T. Sakaki, S. Kominami, K. Inouye, and Y. Yabusaki, "Coexpression of genetically engineered fused enzyme between yeast NADPH-P450 reductase and human cytochrome P450 3A4 and human cytochrome b5 in yeast," *Arch. Biochem. Biophys.*, vol. 381, no. 1, pp. 164–170, 2000.

[342] L. M. Damasceno, C. J. Huang, and C. A. Batt, "Protein secretion in *Pichia pastoris* and advances in protein production," *Appl. Microbiol. Biotechnol.*, vol. 93, no. 1, pp. 31–39, 2012.

[343] R. Aw, G. R. Barton, and D. J. Leak, "Insights into the prevalence and underlying causes of clonal variation through transcriptomic analysis in *Pichia pastoris*," *Appl. Microbiol. Biotechnol.*, vol. 101, no. 12, pp. 5045–5058, Jun. 2017.

[344] C. Schwechheimer and M. J. Kuehn, "Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions," *Nat. Rev. Microbiol.*, vol. 13, pp. 605–619, 2015.

[345] S. Devos *et al.*, "The effect of imipenem and diffusible signaling factors on the secretion of outer membrane vesicles and associated Ax21 proteins in *Stenotrophomonas maltophilia*," *Front. Microbiol.*, vol. 6, no. 298, pp. 1–9, 2015.

[346] S. Devos, S. Stremersch, K. Raemdonck, K. Braeckmans, and B. Devreese, "Intra-and interspecies effects of outer membrane vesicles from *Stenotrophomonas maltophilia* on β-lactam resistance," *Antimicrob. Agents Chemother.*, vol. 60, no. 4, pp. 2516–2518, 2016.

[347] S. Devos *et al.*, "Membrane vesicle secretion and prophage induction in multidrug-resistant *Stenotrophomonas maltophilia* in response to ciprofloxacin stress," *Environ. Microbiol.*, vol. 19, no. 10, pp. 3930–3937, 2017.

[348] Y. M. D. Gnopo, H. C. Watkins, T. C. Stevenson, M. P. DeLisa, and D. Putnam, "Designer outer membrane vesicles as immunomodulatory systems – Reprogramming bacteria for vaccine delivery," *Adv. Drug Deliv. Rev.*, vol. 114, pp. 132–142, 2017.

[349] M. Park, Q. Sun, F. Liu, M. P. Delisa, and W. Chen, "Positional assembly of enzymes on bacterial outer membrane vesicles for cascade reactions," *PLoS One*, vol. 9, no. 5, pp.

1–6, 2014.
[350] J. L. Baker, L. Chen, J. A. Rosenthal, D. Putnam, and M. P. DeLisa, "Microbial biosynthesis of designer outer membrane vesicles," *Curr. Opin. Biotechnol.*, vol. 29, pp. 76–84, 2014.
[351] A. Kulp and M. J. Kuehn, "Biological functions and biogenesis of secreted bacterial outer membrane vesicles," *Annu. Rev. Microbiol.*, vol. 64, pp. 163–184, 2010.
[352] M. Toyofuku, N. Nomura, and L. Eberl, "Types and origins of bacterial membrane vesicles," *Nat. Rev. Microbiol.*, vol. 17, pp. 13–24, 2019.
[353] A. Bernadac, M. Gavioli, J. C. Lazzaroni, S. Raina, and R. Lloubès, "*Escherichia coli* tol-pal mutants form outer membrane vesicles," *J. Bacteriol.*, vol. 180, no. 18, pp. 4872–4878, 1998.
[354] A. J. McBroom, A. P. Johnson, S. Vemulapalli, and M. J. Kuehn, "Outer membrane vesicle production by *Escherichia coli* is independent of membrane instability," *J. Bacteriol.*, vol. 188, no. 15, pp. 5385–5392, 2006.
[355] C. Schwechheimer, D. L. Rodriguez, and M. J. Kuehn, "NlpI-mediated modulation of outer membrane vesicle production through peptidoglycan dynamics in *Escherichia coli*," *Microbiologyopen*, vol. 4, no. 3, pp. 375–389, 2015.
[356] A. J. McBroom and M. J. Kuehn, "Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response," *Mol. Microbiol.*, vol. 63, no. 2, pp. 545–558, 2007.
[357] A. J. Manning and M. J. Kuehn, "Contribution of bacterial outer membrane vesicles to innate bacterial defense," *BMC Microbiol.*, vol. 11, no. 258, pp. 1–14, 2011.
[358] S. W. Kim *et al.*, "The importance of porins and β-lactamase in outer membrane vesicles on the hydrolysis of β-lactam antibiotics," *Int. J. Mol. Sci.*, vol. 21, no. 2822, pp. 1–14, 2020.
[359] T. Baba *et al.*, "Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection," *Mol. Syst. Biol.*, vol. 2, no. 1, pp. 1–11, 2006.
[360] F. W. Studier, "Stable cultures and auto-induction for inducible protein production in *E. coli*," *Struct. Genomics Methods Protoc.*, vol. 1091, pp. 17–32, 2014.
[361] M. Shimazu, A. Mulchandani, and W. Chen, "Cell surface display of organophosphorus hydrolase using ice nucleation protein," *Biotechnol. Prog.*, vol. 17, no. 1, pp. 76–80, 2001.
[362] S. Sandrini, R. Haigh, and P. Freestone, "Fractionation by ultracentrifugation of Gram negative cytoplasmic and membrane proteins," *BIO-PROTOCOL*, vol. 4, no. 21, 2014.
[363] B. van de Waterbeemd *et al.*, "Improved OMV vaccine against *Neisseria meningitidis* using genetically engineered strains and a detergent-free purification process," *Vaccine*, vol. 28, pp. 4810–4816, 2010.
[364] U. Schwaneberg, C. Schmidt-Dannert, J. Schmitt, and R. D. Schmid, "A continuous spectrophotometric assay for P450 BM-3, a fatty acid hydroxylating enzyme, and its mutant F87A," *Anal. Biochem.*, vol. 269, pp. 359–366, 1999.
[365] K. Hantke, "Regulation of ferric iron transport in *Escherichia coli* K12: Isolation of a constitutive mutant," *Mol. Gen. Genet.*, vol. 182, pp. 288–292, 1981.
[366] J. F. Rontani and C. Aubert, "Trimethylsilyl transfer during electron ionization mass spectral fragmentation of some ω-hydroxycarboxylic and ω-dicarboxylic acid trimethylsilyl derivatives and the effect of chain length," *Rapid Commun. Mass Spectrom.*, vol. 18, pp. 1889–1895, 2004.
[367] A. J. Kulp *et al.*, "Genome-wide assessment of outer membrane vesicle production in *Escherichia coli*," *PLoS One*, vol. 10, no. 9, pp. 1–16, 2015.
[368] D. J. Chen *et al.*, "Delivery of foreign antigens by engineered outer membrane vesicle vaccines," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 7, pp. 3099–3104, 2010.

[369] J. Y. Kim *et al.*, "Engineered bacterial outer membrane vesicles with enhanced functionality," *J. Mol. Biol.*, vol. 380, pp. 51–66, 2008.
[370] F. W. Studier, "Protein production by auto-induction in high-density shaking cultures," *Protein Eng. Des. Sel.*, vol. 41, pp. 207–234, 2005.
[371] H. Chutkan, I. MacDonald, A. Manning, and M. J. Kuehn, "Quantitative and qualitative preparations of bacterial outer membrane vesicles," in *Bacterial Cell Surfaces: Methods and Protocols, Methods in Molecular Biology*, vol. 966, A. H. Delcour, Ed. Humana Press Inc., Totowa, NJ, 2013, pp. 259–272.
[372] J. B. Johnston, H. Ouellet, L. M. Podust, and P. R. Ortiz De Montellano, "Structural control of cytochrome P450-catalyzed ω-hydroxylation," *Arch. Biochem. Biophys.*, vol. 507, pp. 86–94, 2011.
[373] D. Kim, M. J. Cryle, J. J. De Voss, and P. R. Ortiz de Montellano, "Functional expression and characterization of cytochrome P450 52A21 from *Candida albicans*," *Arch. Biochem. Biophys.*, vol. 464, pp. 213–220, 2007.
[374] V. O. Popov and V. S. Lamzin, "NAD+-dependent formate dehydrogenase," *Biochem. J.*, vol. 301, pp. 625–643, 1994.
[375] V. I. Tishkov and V. O. Popov, "Catalytic mechanism and application of formate dehydrogenase," *Biochem.*, vol. 69, no. 11, pp. 1252–1267, 2004.
[376] A. S. Bommarius, M. Schwarm, K. Stingl, M. Kottenhahn, K. Huthmacher, and K. Drauz, "Synthesis and use of enantiomerically pure tert-leucine," *Tetrahedron: Asymmetry*, vol. 6, no. 12, pp. 2851–2888, 1995.
[377] K. Hoelsch, I. Sührer, M. Heusel, and D. Weuster-Botz, "Engineering of formate dehydrogenase: Synergistic effect of mutations affecting cofactor specificity and chemical stability," *Appl. Microbiol. Biotechnol.*, vol. 97, pp. 2473–2481, 2013.
[378] V. I. Tishkov and V. O. Popov, "Protein engineering of formate dehydrogenase," *Biomol. Eng.*, vol. 23, pp. 89–110, 2006.
[379] V. I. Tishkov *et al.*, "Rational design of practically important enzymes," *Moscow Univ. Chem. Bull.*, vol. 73, no. 1, pp. 1–6, 2018.
[380] W. Hummel and H. Gröger, "Strategies for regeneration of nicotinamide coenzymes emphasizing self-sufficient closed-loop recycling systems," *J. Biotechnol.*, vol. 191, pp. 22–31, 2014.
[381] S. Kara, J. H. Schrittwieser, and F. Hollmann, "Strategies for cofactor regeneration in biocatalyzed reductions," in *Synthetic Methods for Biologically Active Molecules*, E. Brenna, Ed. Wiley-VCH Verlag GmbH & Co. KGaA, 2013, pp. 209–238.
[382] R. Hatrongjit and K. Packdibamrung, "A novel NADP+-dependent formate dehydrogenase from *Burkholderia stabilis* 15516: Screening, purification and characterization," *Enzyme Microb. Technol.*, vol. 46, pp. 557–561, 2010.
[383] S. Alpdağtaş, S. Yücel, H. A. Kapkaç, S. Liu, and B. Binay, "Discovery of an acidic, thermostable and highly NADP+ dependent formate dehydrogenase from *Lactobacillus buchneri* NRRL B-30929," *Biotechnol. Lett.*, vol. 40, pp. 1135–1147, 2018.
[384] S. Fogal, E. Beneventi, L. Cendron, and E. Bergantino, "Structural basis for double cofactor specificity in a new formate dehydrogenase from the acidobacterium *Granulicella mallensis* MP5ACTX8," *Appl. Microbiol. Biotechnol.*, vol. 99, pp. 9541–9554, 2015.
[385] M. Takacs, O. V. Makhlynets, P. L. Tolbert, and I. V. Korendovych, "Secretion of functional formate dehydrogenase in *Pichia pastoris*," *Protein Eng. Des. Sel.*, vol. 30, no. 5, pp. 381–386, 2017.
[386] Z. E. Duman, B. B. Duraksoy, F. Aktaş, J. M. Woodley, and B. Binay, "High-level heterologous expression of active *Chaetomium thermophilum* FDH in *Pichia pastoris*," *Enzyme Microb. Technol.*, vol. 137, no. 109552, pp. 1–5, 2020.

[387] A. Liu, R. Feng, and B. Liang, “Microbial surface displaying formate dehydrogenase and its application in optical detection of formate,” *Enzyme Microb. Technol.*, vol. 91, pp. 59–65, 2016.
[388] N. K. Khatri and F. Hoffmann, “Impact of methanol concentration on secreted protein production in oxygen-limited cultures of recombinant *Pichia pastoris*,” *Biotechnol. Bioeng.*, vol. 93, no. 5, pp. 871–879, 2006.
[389] H. W. Jiang, Q. Chen, J. Pan, G. W. Zheng, and J. H. Xu, “Rational engineering of formate dehydrogenase substrate/cofactor affinity for better performance in NADPH regeneration,” *Appl. Biochem. Biotechnol.*, vol. 192, pp. 530–543, 2020.
[390] M. H. Hsu, B. R. Baer, A. E. Rettie, and E. F. Johnson, “The crystal structure of cytochrome P450 4B1 (CYP4B1) monooxygenase complexed with octane discloses several structural adaptations for ω-hydroxylation,” *J. Biol. Chem.*, vol. 292, no. 13, pp. 5610–5621, 2017.
[391] S. M. Hoffmann, H. R. Danesh-Azari, C. Spandolf, M. J. Weissenborn, G. Grogan, and B. Hauer, “Structure-guided redesign of CYP153AM.aq for the improved terminal hydroxylation of fatty acids,” *ChemCatChem*, vol. 8, p. 3178, 2016.
[392] Z. Li, Y. Jiang, X. F. P. Guengerich, L. Ma, S. Li, and W. Zhang, “Engineering cytochrome P450 enzyme systems for biomedical and biotechnological applications,” *J. Biol. Chem.*, vol. 295, no. 3, pp. 833–849, 2020.
[393] J. L. Kadurugamuwa and T. J. Beveridge, “Membrane vesicles derived from *Pseudomonas aeruginosa* and *Shigella flexneri* can be integrated into the surfaces of other Gram-negative bacteria,” *Microbiology*, vol. 145, pp. 2051–2060, 1999.
[394] D. M. B. Post, D. S. Zhang, J. S. Eastvold, A. Teghanemt, B. W. Gibson, and J. P. Weiss, “Biochemical and functional characterization of membrane blebs purified from *Neisseria meningitidis* serogroup B,” *J. Biol. Chem.*, vol. 280, no. 46, pp. 38383–38394, 2005.
[395] N. J. Alves, K. B. Turner, K. A. DiVito, M. A. Daniele, and S. A. Walper, “Affinity purification of bacterial outer membrane vesicles (OMVs) utilizing a His-tag mutant,” *Res. Microbiol.*, vol. 168, pp. 139–146, 2017.
[396] S.-K. Yim, H.-C. Jung, C.-H. Yun, and J.-G. Pan, “Functional expression in *Bacillus subtilis* of mammalian NADPH-cytochrome P450 oxidoreductase and its spore-display,” *Protein Expr. Purif.*, vol. 63, pp. 5–11, 2009.

Curriculum vitae

ir. Delphine Devriese

delphinedevriese@gmail.com
https://www.linkedin.com/in/delphine-devriese-5b9775102/

Education

2017-2021	Doctor of Science in Biochemistry & Biotechnology Ghent University Laboratory of Microbiology – Protein Research Unit (LM-UGent PRU) → "A CYP solution: an enzyme-based process for the production of long-chain α,ω-dicarboxylic acids"
2014-2016	Master of Bioscience Engineering – Cell and Gene Biotechnology Ghent University → "Tumor cell exosome biogenesis and the actin cytoskeleton: exploring the interconnection using cortactin nanobodies"
2011-2014	Bachelor of Bioscience Engineering – Cell and Gene Biotechnology Ghent University

Skills

Technical skills

- Molecular cloning
- Recombinant protein production
  - *Pichia pastoris*
  - *Escherichia coli*
- ÄKTA protein purification
  - IMAC
  - SEC
- Differential and density ultracentrifugation
- SDS-PAGE and western blot
- MALDI-TOF MS
- Spectrophotometric assays
- GC-MS analysis
- Human cell culture
- Fluoresence microscopy

Computer skills

- Good knowledge of Microsoft Office
- Knowledge of Matlab and R
- Basic knowledge of Python

Languages

- Dutch: native language
- English: professional
- French: basic

Courses and conferences

Specialist courses

2019	Gentle hands-on introduction to Python programming
2019	Protein Purification
2018	EMBO workshop “Enzymes, biocatalysis and chemical biology: The new frontiers” (poster presentation)

Transferable skills

2020	Knowledge for growth (poster presentation)
2018	Effective Graphical Display
2017	Advanced Academic English: Conference Skills – Academic Posters

Conferences

2019	BioTrans (poster presentation)

Guidance of students and practical courses

2019	Integrated practical course Biochemistry & Microbiology: 2D-PAGE
2018	Integrated practical course Biochemistry & Microbiology: 2D-PAGE
	Guidance of a master 1 student
2017	Integrated practical course Biochemistry & Microbiology: 2D-PAGE

Publications

E. Beghein, D. Devriese, E. Van Hoey, and J. Gettemans, “Cortactin and fascin-1 regulate extracellular vesicle release by controlling endosomal trafficking or invadopodia formation and function,” *Sci. Rep.*, vol. 8, no. 1, pp. 1–16, 2018.

In preparation:
“Display of the self-sufficient CYP102A1 on the surface of *E. coli*-derived OMVs”
“Recombinant production of class II CYP enzymes” (review)

Addendum

Figure A1: Expression vectors for the production of CPRtr and CYPtr (left and right, respectively) in *Pichia pastoris* and secretion to the medium. The amino acid sequence of the artificial N-tag is MGSSHHHHHHSSGLVPRGSHMLDIM and contains a Histag and a thrombin cleavage site.

Figure A2: Expression vectors for the production of CYPnt and CPRnt in the microsomal membranes of *Pichia pastoris*.

Figure A3: SDS-PAGE analysis of microsomes after CYPnt production in *P. pastoris*. A: Comparison between microsomes after enzymatic cell lysis versus mechanical cell lysis. B: Microsomal protein obtained after enzymatic cell lysis. The lysate was centrifuged at 20 000 g or 30 000 g or no intermediate centrifugation step was included before collecting the microsome from the lysate at 100 000 g.

Figure A4: SDS-PAGE analysis of microsomes after a 24 h, 48 h, 72 h or 96 h CYPnt expression phase. The blue arrow indicates a band coinciding with the expected MW of CYPnt. MALDI-TOF MS identified this band as 60S ribosome subunit biogenesis protein and not CYPnt.

Figure A5: Expression vector encoding the surface display anchor InaK in a truncated form (InaKtr).

Figure A6: Peptide mass fingerprint of gel band appearing in between 100 kDa and 150 kDa in the medium sample, collected after InaKtrBM3 production in JC8031 for 3 h in TB medium. BM3 was identified in this gel band and no peptides, specific for the InaKtr anchor were found.

Figure A7: Peptide mass fingerprint of gel band appearing in 150 kDa in the OM sample, collected after InaKtrBM3 production in JC8031 for 3 h in TB medium. InaKtrBM3 was identified in this gel band. Peptides, specific for the InaKtr anchor were found and confirmed to be derived from InaKtrBM3 by MS/MS. These peptides are indicated with the grey arrows.

Figure A8: Expression vector encoding the formate dehydrogenase from *Burkholderia stabilis*. Left: Expression vector for the production and secretion in the recombinant host *Pichia pastoris*. Right: Expression vector for the production in the recombinant host *Escherichia coli* and translocation to the periplasm.

Figure A9: Alignment of FDH from *Burkholderia stabilis* to the extensively investigated FDH from *Candida boidinii* and *Pseudomonas sp. 101*.

Table A1: Codon optimized sequences encoding truncated CPR-a (*cprtr*), truncated CYP52A13 (*cyptr*), nontruncated CYP52A13 (*cypnt*) and nontruncated CPR-a (*cprnt*). UniProtKB entry CPR-a: Q66T17, UniProtKB entry CYP52A13: Q874J4.

Table A2: Lists of transitions selected for MRM validation

Table A3: Codon optimized sequences encoding the truncated ice nucleation protein InaK (*inaKtr*). UniProtKB entry InaK from *Pseudomonas syringae*: O30611.

Table A4: Codon optimized sequences encoding formate dehydrogenase (*fdh*) from *Burkholderia stabilis* 15516 for heterologous expression in *Pichia pastoris* or *Escherichia coli*, respectively. UniProtKB entry FDH from *B. stabilis* 15516: B5A8W5.

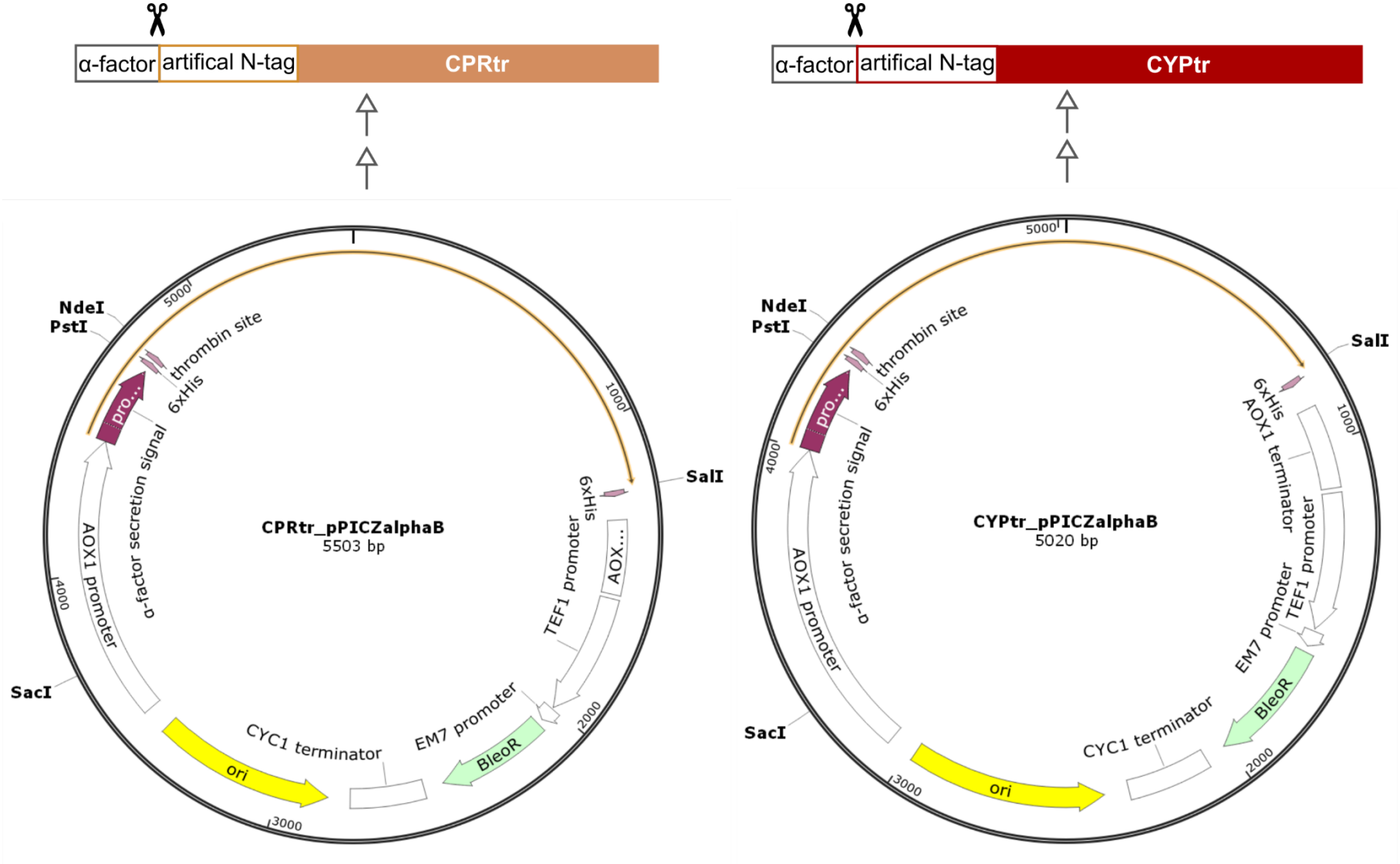

Figure A1: Expression vectors for the production of CPRtr and CYPtr (left and right, respectively) in *Pichia pastoris* and secretion to the medium. The amino acid sequence of the artificial N-tag is MGSSHHHHHHSSGLVPRGSHMLDIM and contains a Histag and a thrombin cleavage site.

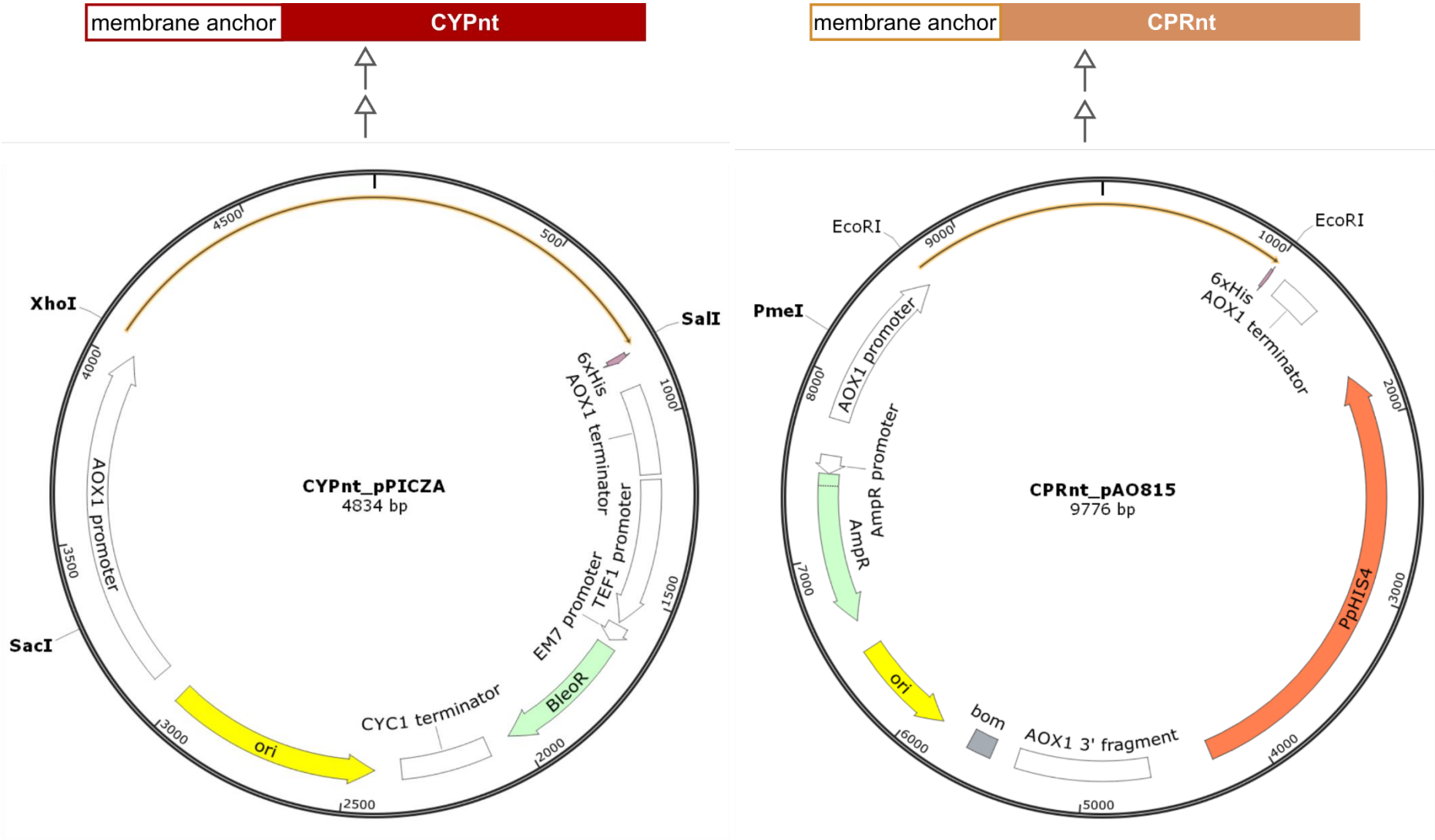

Figure A2: Expression vectors for the production of CYPnt and CPRnt in the microsomal membranes of *Pichia pastoris*.

Table A1: Codon optimized sequences encoding truncated CPR-a (*cprtr*), truncated CYP52A13 (*cyptr*), nontruncated CYP52A13 (*cypnt*) and nontruncated CPR-a (*cprnt*). UniProtKB entry CPR-a: Q66T17, UniProtKB entry CYP52A13: Q874J4.

Gene	Codon optimized sequence
cprtr	**CTGCAGCT** ATGGGTTCTTCACACCACCACCACCACCACAGTTCAGGTCTTGTCCCTAGAGGTTCACATATGCTTGATATTATGTCCCTTAACACCGATTCTGGT TCTAACTCTAGAGATGTTTTGTCTACTTTGAAGAAAAACAATAAGAACACTTTGTTGTTGTTCGGTTCTCAAACTGGTACTGCTGAGGATTACGCT AACAAGTTGTCTAGAGAATTGCATTCTAGATTCGGTTTGAAGACTATGGTTGCTGATTTCGCTGATTACGATTGGGATAACTTCGGAGATATCAC TGAGGATATCTTGGTTTTCTTTATCGTTGCTACTTATGGTGAAGGTGAACCTACTGATAATGCTGATGAATTCCACACTTGGTTGACTGAAGAGG CTGATACTTTGTCTACTTTGAAGTACACTGTTTTTGGTTTGGGTAACTCTACTTATGAGTTTTTCAACGCTATCGGTAGAAAGTTCGATAGATTGT TGTCTGAAAAGGGTGGAGATAGATTTGCTGAGTACGCTGAAGGAGATGATGGTACTGGTACTTTGGATGAGGATTTCATGGCTTGGAAGGATAA CGTTTTCGATGCTTTGAAAAACGATTTGAACTTCGAAGAGAAGGAGTTGAAGTACGAACCTAACGTTAAGTTGACTGAAAGAGATGATTTGTCT GCTGCTGATTCTCAAGTTTCTTTGGGAGAGCCAAACAAGAAATACATCAACTCTGAAGGTATTGATTTGACTAAAGGTCCTTTTGATCATACTCA CCCATATTTGGCTAGAATTACTGAGACTAGAGAATTGTTCTCTTCTAAGGATAGACATTGTATTCACGTTGAGTTCGATATCTCTGAATCTAACTT GAAGTACACTACTGGAGATCATTTGGCTATTTGGCCATCTAACTCTGATGAAAACATCAAGCAATTCGCTAAGTGTTTTGGTTTGGAGGATAAGT TGGATACTGTTATTGAATTGAAGGCTTTGGATTCTACTTACACTATTCCATTTCCTACTCCAATTACTTATGGTGCTGTTATTAGACATCACTTGG AAATTTCTGGTCCTGTTTCTAGACAATTTTTCTTGTCTATTGCTGGTTTCGCTCCAGATGAAGAGACTAAGAAAGCTTTTACTAGATTGGGTGGAG ATAAGCAAGAATTCGCTGCTAAGGTTACTAGAAGAAAGTTTAACATTGCTGATGCTTTGTTGTATTCTTCTAACAATGCTCCTTGGTCTGATGTTC CATTTGAGTTCTTGATTGAAAACGTTCCTCACTTGACTCCAAGATACTACTCTATCTCTTCTTCTTTGTCTGAGAAGCAATTGATTAATGTTAC TGCTGTTGTTGAAGCTGAAGAGGAAGCTGATGGTAGACCTGTTACTGGTGTTGTTACTAACTTGTTGAAGAATGTTGAGATTGTTCAAAACAAGA CTGGTGAAAAACCTTTGGTTCATTACGATTTGTCTGGTCCACGTGGAAAGTTTAACAAGTTTAAGTTGCCAGTTCACGTTAGAAGATCTAACTTC AAGTTGCCTAAAAATTCTACTACTCCAGTTATTTTGATTGGTCCTGGTACTGGTGTTGCTCCATTGAGAGGTTTTGTTAGAGAGAGAGTTCAACA AGTTAAGAACGGTGTTAACGTTGGTAAAACTTTGTTGTTCTATGGTTGTAGAAACTCTAACGAAGATTTCTTGTACAAGCAAGAGTGGGCTGAAT ATGCTTCTGTTTTGGGAGAGAACTTCGAAATGTTCAACGCTTTTTCTAGACAAGATCCATCTAAGAAAGTTTACGTTCAAGATAAGATCTTGGAA AATTCTCAATTGGTTCATGAGTTGTTGACTGAAGGTGCTATTATCTACGTTTGTGGAGATGCTTCTAGAATGGCTAGAGATGTTCAAACTACTATC TCTAAGATCGTTGCTAAGTCTAGAGAGATTTCTGAAGATAAGGCAGCCGAGTTGGTCAAGAGTTGGAAGGTTCAGAATAGATATCAGGAAGATG TTTGGTAG **GTCGAC**
cyptr	**CTGCAGCT** ATGGGATCTTCCCATCACCATCACCATCACTCAAGTGGATTGGTTCCAAGAGGTTCTCATATGTTGGATATTAAACTTGGTGCAAAGCCTGTTAC TCAAAAGCAGACCGACGGATGTACAGGTATCAAAGCTCCATTGGAATTGCTTAAGAAAAAGAGTGATGGTACTTTGATTGACTTTACCCTTCAG AGAATCCATGATTTGGACAGACCAGATATTCCTACCTTTACATTCCCAGTTTTCTCTATCAACTTGGTCAATACCCTTGAACCTGAGAACATTAAA GCCATCTTGGCAACTCAATTCAACGATTTCAGTTTGGGTACCAGACATTCTCACTTTGCTCCATTGCTTGGAGATGGTATTTTCACTTTGGACGGA GCCGGTTGGAAACACTCTAGATCCATGTTGAGACCACAATTTGCTAGAGAGCAGATCTCCCATGTTAAGTTGCTTGAACCTCACGTTCAAGTCTT TTTCAAACATGTTAGAAAGGCTCAGGGTAAAACTTTCGATATCCAAGAATTGTTTTTCAGACTTACAGTTGACTCTGCCACTGAGTTTTTGTTCGG AGAATCAGTCGAGAGTCTTAGAGATGAAAGTATTGGAATGTCTATCAACGCCTTGGATTTTGACGGAAAAGCAGGTTTTGCAGACGCTTTCAAC TACTCCCAAAATTATTTGGCTTCAAGAGCCGTTATGCAACAGTTGTACTGGGTCCTTAACGGTAAAAAGTTTAAGGAATGCAATGCCAAAGTTCA TAAGTTCGCAGATTACTATGTCAACAAAGCTTTGGACCTTACTCCAGAACAATTGGAGAAGCAGGATGGTTACGTTTTCTTGTATGAGCTTGTCA AGCAAACCAGAGATAAACAGGTTTTGAGAGACCAATTGCTTAATATTATGGTCGCTGGAAGAGACACTACCGCCGGTTTGCTTTCCTTCGTTTTC

	TTTGAGTTGGCTAGAAACCCTGAAGTCACTAATAAGTTGAGAGAAGAGATCGAAGATAAATTTGGATTGGGTGAAAACGCCTCAGTTGAGGACA TCTCTTTCGAATCCTTGAAGTCTTGTGAGTACTTGAAAGCAGTTCTTAATGAAACCTTGAGACTTTATCCATCCGTTCCTCAGAACTTTAGAGTCG CTACAAAGAATACAACTTTGCCAAGAGGTGGAGGTAAAGATGGATTGTCACCTGTTCTTGTCAGAAAGGGACAAACTGTTATCTACGGTGTCTA TGCTGCCCATAGAAACCCAGCAGTTTACGGTAAAGACGCTTTGGAATTCAGACCAGAGAGATGGTTCGAACCTGAGACTAAAAAGTTGGGATGG GCTTTTCTTCCATTCAATGGAGGTCCTAGAATCTGCTTGGGTCAACAGTTTGCACTTACCGAGGCTAGTTACGTTACAGTCAGATTGCTTCAAGA ATTCGCACACTTGTCTATGGATCCTGACACAGAGTATCCACCTAAAAAGATGTCCCATTTGACTATGTCACTTTTCGATGGTGCCAATATTGAAA TGTATTAA **GTCGAC**
cypnt	**CTCGAG** ATGACCGTCCACGACATCATCGCTACTTACTTTACAAAATGGTACGTTATTGTTCCACTTGCCCTTATCGCCTACAGAGTCCTTGACTACTTCTAC GGTAGATACTTGATGTACAAGTTGGGTGCTAAGCCATTTTTCCAAAAGCAAACTGATGGTTGTTTTGGTTTCAAAGCTCCTTTGGAATTGTTGAA GAAAAAGTCTGATGGTACTTTGATCGATTTCACTTTGCAAAGAATCCACGATTTGGATAGACCAGATATTCCTACTTTTACTTTCCCAGTTTTCTC TATTAACTTGGTTAACACTTTGGAACCTGAGAACATTAAAGCTATTTTGGCTACTCAATTCAACGATTTCTCTTTGGGTACTAGACATTCTCACTT TGCTCCATTGTTGGGAGATGGTATTTTCACTTTGGATGGTGCTGGTTGGAAACATTCTAGATCTATGCTTAGACCACAATTTGCTAGAGAGCAAA TTTCTCATGTTAAGTTGTTGGAACCTCACGTTCAAGTTTTCTTTAAGCATGTTAGAAAGGCTCAAGGTAAAACTTTCGATATCCAAGAATTGTTTT TCAGATTGACTGTTGATTCTGCTACTGAGTTTTTGTTCGGTGAATCTGTTGAGTCTTTGAGAGATGAATCTATTGGAATGTCTATTAACGCTTTGG ATTTTGATGGTAAAGCTGGTTTTGCTGATGCTTTCAACTACTCTCAAAACTATTTGGCTTCTAGAGCTGTTATGCAACAATTGTACTGGGTTTTGA ACGGTAAAAAGTTTAAGGAATGTAACGCTAAGGTTCACAAGTTCGCTGATTACTACGTTAACAAAGCTTTGGATTTGACTCCAGAACAATTGGA GAAGCAAGATGGTTACGTTTTCTTGTATGAGTTGGTTAAGCAAACTAGAGATAAGCAAGTTTTGAGAGATCAATTGTTGAACATCATGGTTGCTG GTAGAGATACTACTGCTGGTTTGTTGTCTTTCGTTTTCTTTGAGTTGGCTAGAAACCCTGAAGTTACTAATAAGTTGAGAGAAGAGATCGAAGAT AAATTTGGTTTGGGTGAAAACGCTTCTGTTGAGGATATTTCTTTCGAATCTTTGAAGTCTTGTGAGTACTTGAAAGCTGTCTTGAACGAAACTTTG AGATTGTACCCATCTGTTCCTCAAAACTTCAGAGTTGCTACTAAGAATACTACTTTGCCAAGAGGTGGTGGTAAAGATGGTTTGTCTCCTGTTTTG GTTAGAAAGGGTCAAACTGTTATCTACGGTGTTTATGCTGCTCACAGAAACCCAGCTGTTTATGGTAAAGATGCTTTGGAATTCAGACCAGAGA GATGGTTCGAACCTGAGACTAAAAAGTTGGGTTGGGCTTTCTTGCCATTCAATGGTGGTCCTAGAATTTGTTTGGGTCAACAATTTGCTTTGACT GAGGCTTCTTACGTTACTGTTAGATTGTTGCAAGAATTCGCTCATTTGTCTATGGATCCTGATACTGAGTATCCACCTAAGAAAATGTCCCACTTG ACTATGAGTTTGTTTGACGGTGCTAATATTGAAATGTACTAA **GTCGAC**
cprnt	**GAATTCACC** ATGGCATTGGATAAGTTGGATTTGTATGTCATCATTACCCTTGTTGTCGCAGTCGCAGCCTATTTTGCTAAGAACCAGTTTTTGGATCAACCTCAA GATACTGGTTTCTTGAACACTGATTCTGGTTCTAATTCTAGAGATGTTTTGTCTACTTTGAAGAAAAACAATAAGAACACTTTGTTGTTGTTCGGT TCTCAAACTGGTACTGCTGAGGATTACGCTAATAAGTTGTCTAGAGAATTGCATTCTAGATTCGGTTTGAAGACTATGGTTGCTGATTTCGCTGA TTACGATTGGGATAACTTCGGAGATATCACTGAGGATATCTTGGTTTTCTTTATCGTTGCTACTTATGGAGAGGGTGAACCAACTGATAATGCTG ATGAATTTCACACTTGGTTGACTGAAGAGGCTGATACTTTGTCTACTTTGAAGTACACTGTTTTTGGTTTGGGTAACTCTACTTACGAGTTTTTCA ACGCTATCGGTAGAAAGTTCGATAGATTGTTGTCTGAAAAGGGTGGAGATAGATTTGCTGAGTACGCTGAAGGAGATGATGGTACTGGTACTTT GGATGAGGATTTCATGGCTTGGAAGGATAACGTTTTCGATGCTTTGAAAAACGATTTGAACTTCGAAGAGAAGGAGTTGAAGTACGAACCTAAC GTTAAGTTGACTGAAAGAGATGATTTGTCTGCTGCTGATTCTCAAGTTTCTTTGGGAGAGCCAAACAAGAAATACATCAACTCTGAAGGTATTGA TTTGACTAAAGGTCCTTTTGATCATACTCACCCATATTTGGCTAGAATTACTGAGACTAGAGAATTGTTCTCTTCTAAGGATAGACATTGTATTCA CGTTGAGTTTGATATCTCTGAATCTAACTTGAAGTACACTACTGGAGATCATTTGGCTATTTGGCCTTCTAACTCTGATGAAAACATCAAGCAATT

CGCTAAGTGTTTTGGTTTGGAGGATAAGTTGGATACTGTTATTGAATTGAAGGCTTTGGATTCTACTTACACTATTCCATTTCCTACTCCAATTAC
TTATGGTGCTGTTATTAGACATCACTTGGAAATTTCTGGTCCTGTTTCTAGACAATTTTTCTTGTCTATTGCTGGTTTCGCTCCAGATGAAGAGACT
AAGAAAGCTTTTACTAGATTGGGTGGAGATAAGCAAGAATTTGCTGCTAAGGTTACTAGAAGAAAGTTTAACATTGCTGATGCTTTGTTGTATTC
TTCTAACAATGCTCCTTGGTCTGATGTTCCATTTGAGTTCTTGATTGAAAACGTTCCTCACTTGACTCCAAGATACTACTCTATCTCTTCTTCTTCT
TTGTCTGAGAAGCAATTGATTAATGTTACTGCTGTTGTTGAAGCTGAAGAGGAAGCTGATGGTAGACCTGTTACTGGTGTTGTTACTAACTTGTT
GAAGAATGTTGAGATTGTTCAAAACAAGACTGGTGAAAAACCTTTGGTTCATTACGATTTGTCTGGTCCACGTGGAAAGTTTAACAAGTTTAAGT
TGCCAGTTCACGTTAGAAGATCCAACTTCAAGTTGCCTAAAAATTCTACTACTCCAGTTATTTTGATTGGTCCTGGTACTGGTGTTGCTCCATTGA
GAGGTTTTGTTAGAGAGAGAGTTCAACAAGTTAAGAACGGTGTTAACGTTGGTAAAACTTTGTTGTTCTATGGTTGTAGAAACTCTAACGAAGAT
TTCTTGTACAAGCAAGAGTGGGCTGAATATGCTTCTGTTTTGGGAGAGAACTTCGAAATGTTCAACGCTTTTTCTAGACAAGATCCATCTAAGAA
AGTTTACGTTCAAGATAAAATTTTGGAAAACTCTCAATTGGTTCATGAGTTGTTGACTGAAGGTGCTATTATCTACGTTTGTGGAGATGCTTCTAG
AATGGCTAGAGATGTTCAAACTACTATCTCTAAGATCGTTGCTAAGTCTAGAGAGATTTCTGAAGATAAGGCTGCTGAATTGGTTAAAAGTTGG
AAGGTTCAGAATAGATACCAGGAGGATGTTTGGCACCACCACCACCACCATTAG
GAATTC

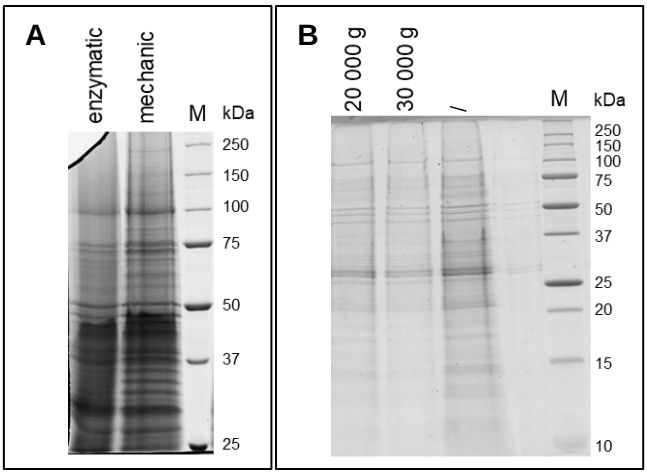

Figure A3: SDS-PAGE analysis of microsomes after CYPnt production in *P. pastoris*. A: Comparison between microsomes after enzymatic cell lysis versus mechanical cell lysis. B: Microsomal protein obtained after enzymatic cell lysis. The lysate was centrifuged at 20 000 g or 30 000 g or no intermediate centrifugation step was included before collecting the microsome from the lysate at 100 000 g.

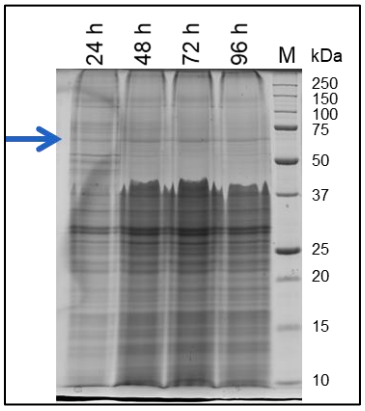

Figure A4: SDS-PAGE analysis of microsomes after a 24 h, 48 h, 72 h or 96 h CYPnt expression phase. The blue arrow indicates a band coinciding with the expected MW of CYPnt. MALDI-TOF MS identified this band as 60S ribosome subunit biogenesis protein and not CYPnt.

Table A2: Lists of transitions selected for MRM validation

Protein	Peptide sequence	Precursor (m/z)	Product ion (m/z)	Collision energy (V)
CYP52A13	SDGTLIDFTLQR	683.354073	1005.572757 892.488693 779.404629	24
	AILATQFNDFSLGTR	827.433386	1184.569462 1056.510885 909.442471	30
	TFDIQELFFR	658.337695	952.525078 839.441014 711.382437	23
	FADYYVNK	510.245275	801.377745 686.350802 523.287474	18
	QDGYVFLYELVK	737.384842	1010.592095 911.523681 764.455267	26
	LYPSVPQNFR	610.824755	944.494841 847.442077 760.410049 661.341635	22
	DGLSPVLVR	478.282191	670.424636 583.392608 486.339844	17
	LGWAFLPFNGGPR	716.380228	1004.531226 857.462812 744.378748	25
	EEGVSLEK	445.726915	632.361367 575.339903 476.271489	15

Protein	Peptide sequence	Precursor (m/z)	Product ion (m/z)	Collision energy (V)
CYP52A13 (continued)	AVLNETLR	458.2665	518.2933 632.3362 745.4203	16
	AGFADAFNYSQNYLASR	947.9396	1101.5323 1215.5753 1362.6437	34
BSA	AEFVEVTK	461.74765	722.408317 575.339903 476.271489	16
	QTALVELLK	507.813324	785.513117 714.476003 601.391939	18

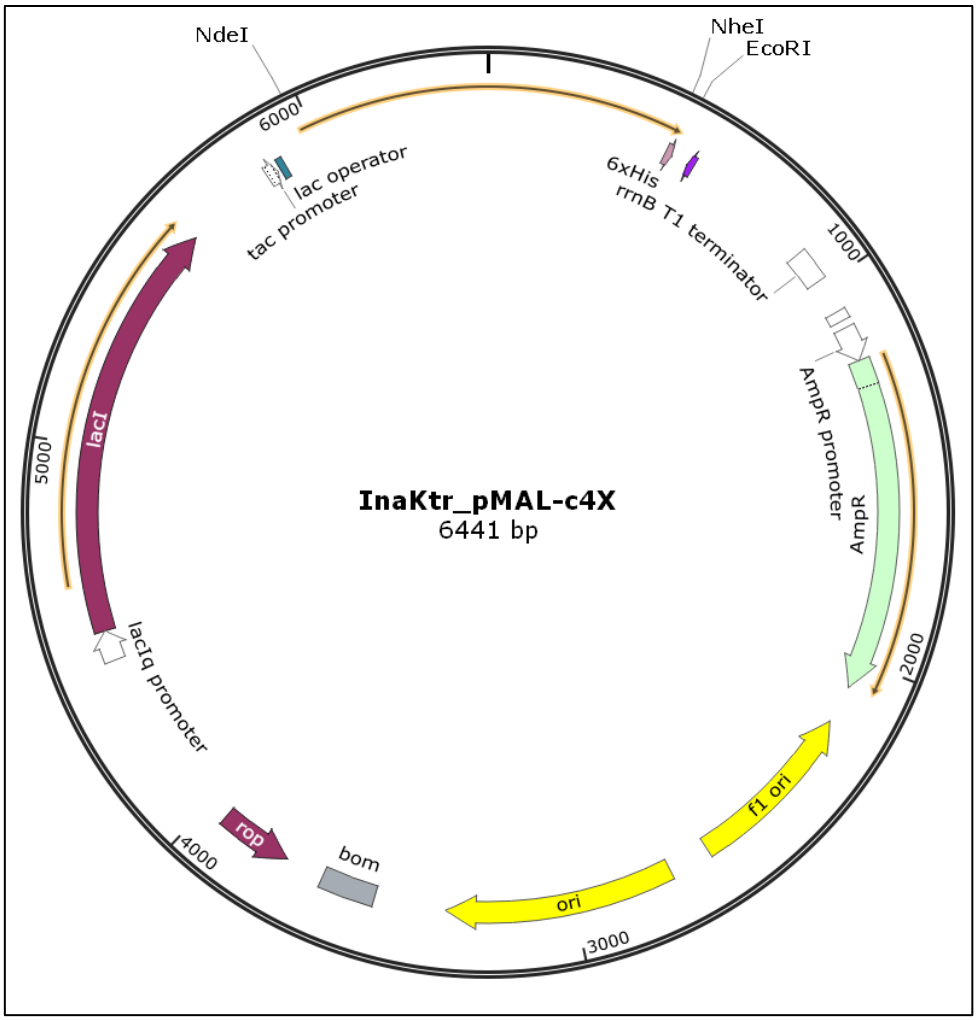

Figure A5: Expression vector encoding the surface display anchor InaK in a truncated form (InaKtr).

Table A3: Codon optimized sequences encoding the truncated ice nucleation protein InaK (*inaKtr*). UniProtKB entry InaK from *Pseudomonas syringae*: O30611.

Gene	Codon optimized sequence
inaKtr	**CAT<u>ATG</u>** ACCCTGGATAAGGCGCTGGTTCTGCGTACCTGCGCGAATAACATGGCGGATCACTGC GGTCTGATTTGGCCGGCGAGCGGCACCGTGGAGAGCCGTTACTGGCAGAGCACCCGT CGTCACGAAAACGGTCTGGTTGGCCTGCTGTGGGGTGCGGGTACCAGCGCGTTCCTG AGCGTTCATGCGGATGCGCGTTGGATCGTGTGCGAGGTGGCGGTTGCGGATATCATT AGCCTGGAGGAACCGGGCATGGTTAAGTTCCCGCGTGCGGAAGTGGTTCACGTGGGT GACCGTATCAGCGCGAGCCACTTTATTAGCGCGCGTCAAGCGGATCCGGCGAGCACC AGCACCAGCACCAGCACCAGCACCCTGACCCCGATGCCGACCGCGATTCCGACCCCG ATGCCGGCGGTGGCGAGCGTTACCCTGCCGGTTGCGGAGCAAGCGCGTCACGAAGTG TTTGATGTTGCGAGCGTGAGCGCGGCGGCGGCGCCGGTTAACACCCTGCCGGTGACC ACCCCGCAGAACCTGCAAACCGCGACCTACGGTAGCACCCTGAGCGGTGATAACCAC AGCCGTCTGATCGCGGGTTATGGCAGCAACGAAACCGCGGGTAACCACAGCGATCTG ATTACCGGTGGCCATGACTGCACCCTGATGGCGGGTGATCAGAGCCGTCTGACCGCG GGCAAGAACAGCGTGCTGACCGCGGGTGCGCGTAGCAAACTGATTGGTAGCGAAGG CAGCACCCTGAGCGCGGGTGAAGACAGCACCCTGATTTTCCGTCTGTGGGATGGCAA GCGTTATCGTCAACTGGTTGCGCGTACCGGCGAGAACGGCGTGGAAGCGGACATCCC GTACTATGTTAACGAGGACGATGACATTGTGGATAAACCGGACGAAGACGACGACT GGATTGAAGTTAAGGACCCGGGCGCTAGCCATCATCACCATCATCACTAA **GAATTC**

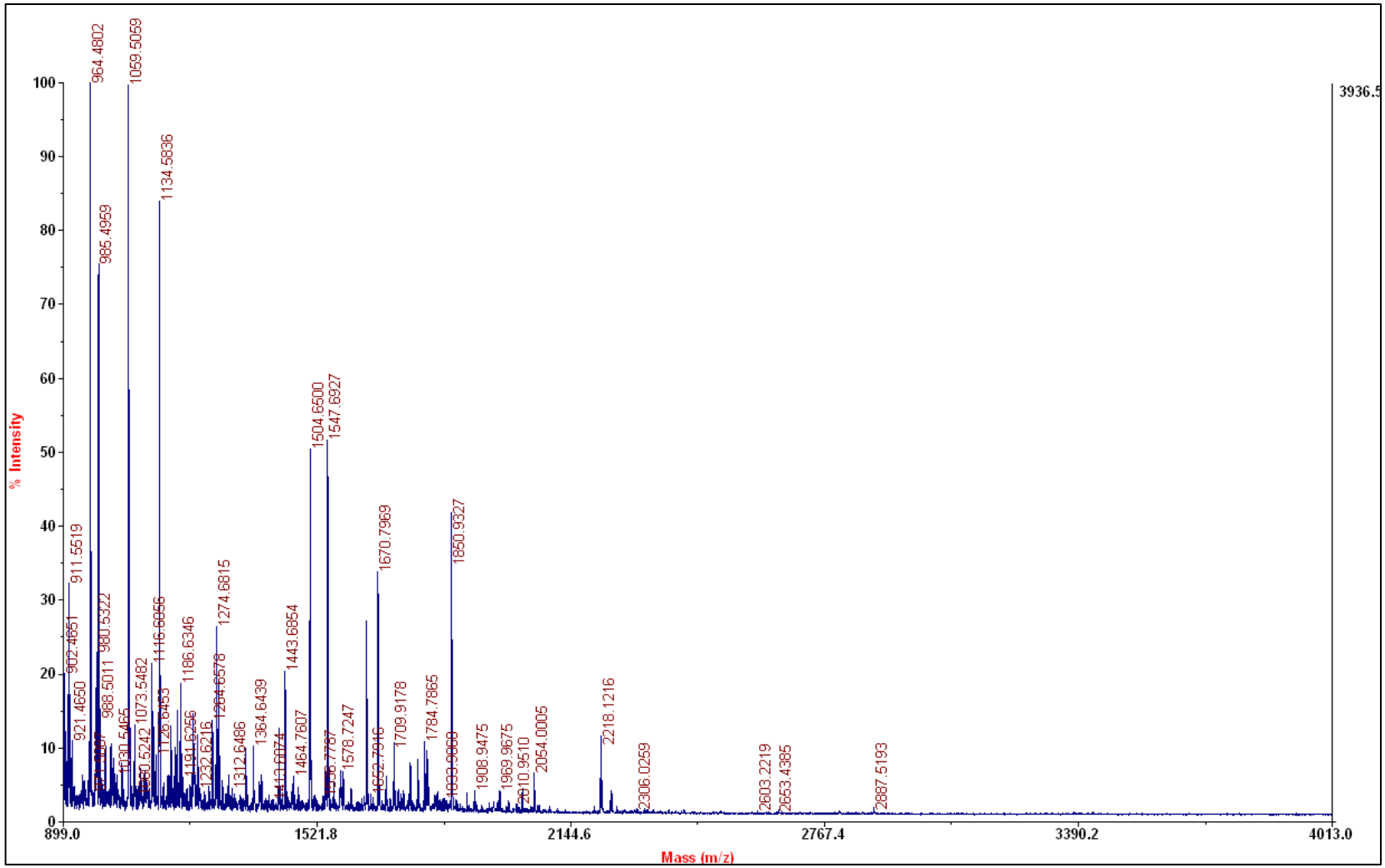

Figure A6: Peptide mass fingerprint of gel band appearing in between 100 kDa and 150 kDa in the medium sample, collected after InaKtrBM3 production in JC8031 for 3 h in TB medium. BM3 was identified in this gel band and no peptides, specific for the InaKtr anchor were found.

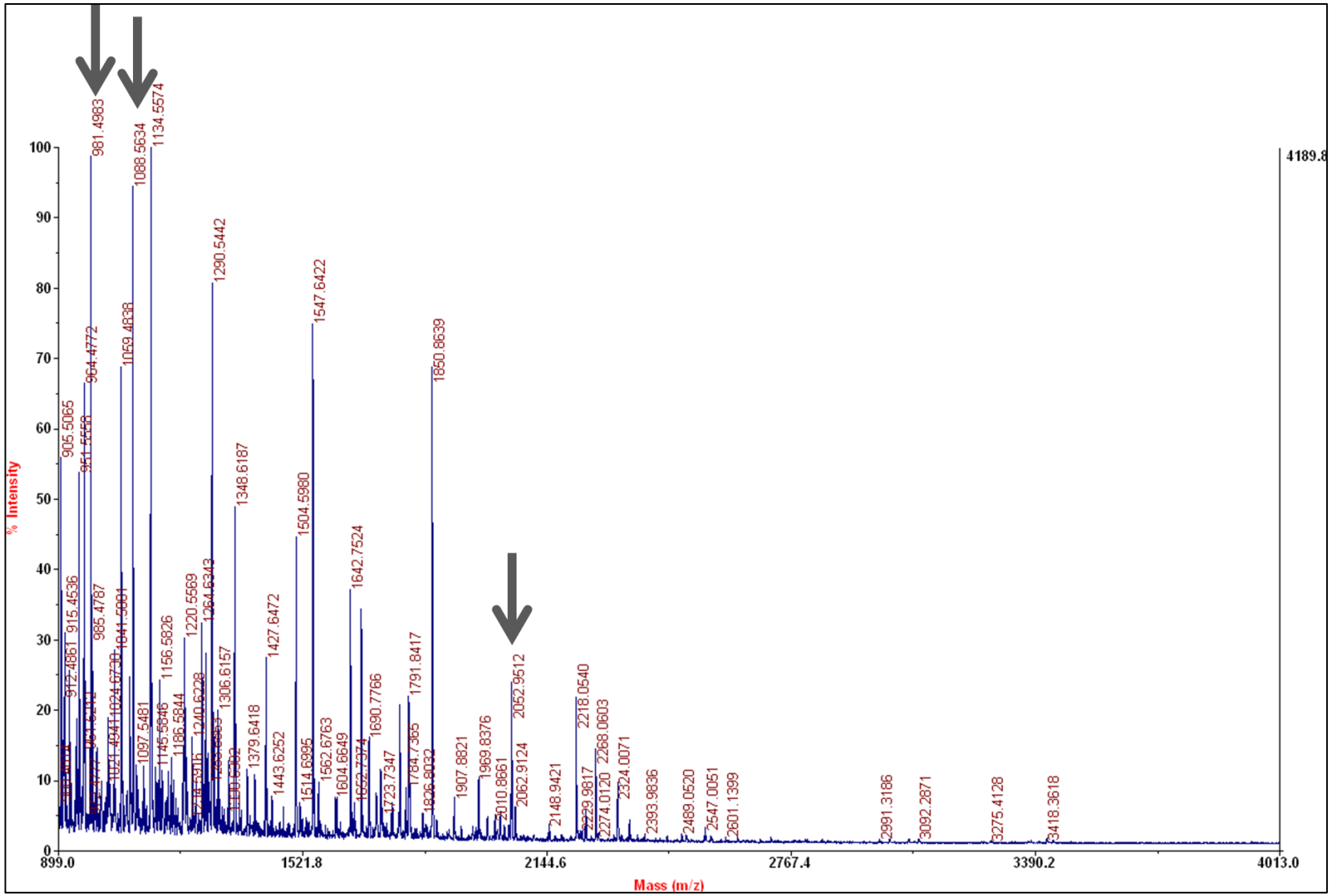

Figure A7: Peptide mass fingerprint of gel band appearing in 150 kDa in the OM sample, collected after InaKtrBM3 production in JC8031 for 3 h in TB medium. InaKtrBM3 was identified in this gel band. Peptides, specific for the InaKtr anchor were found and confirmed to be derived from InaKtrBM3 by MS/MS. These peptides are indicated with the grey arrows.

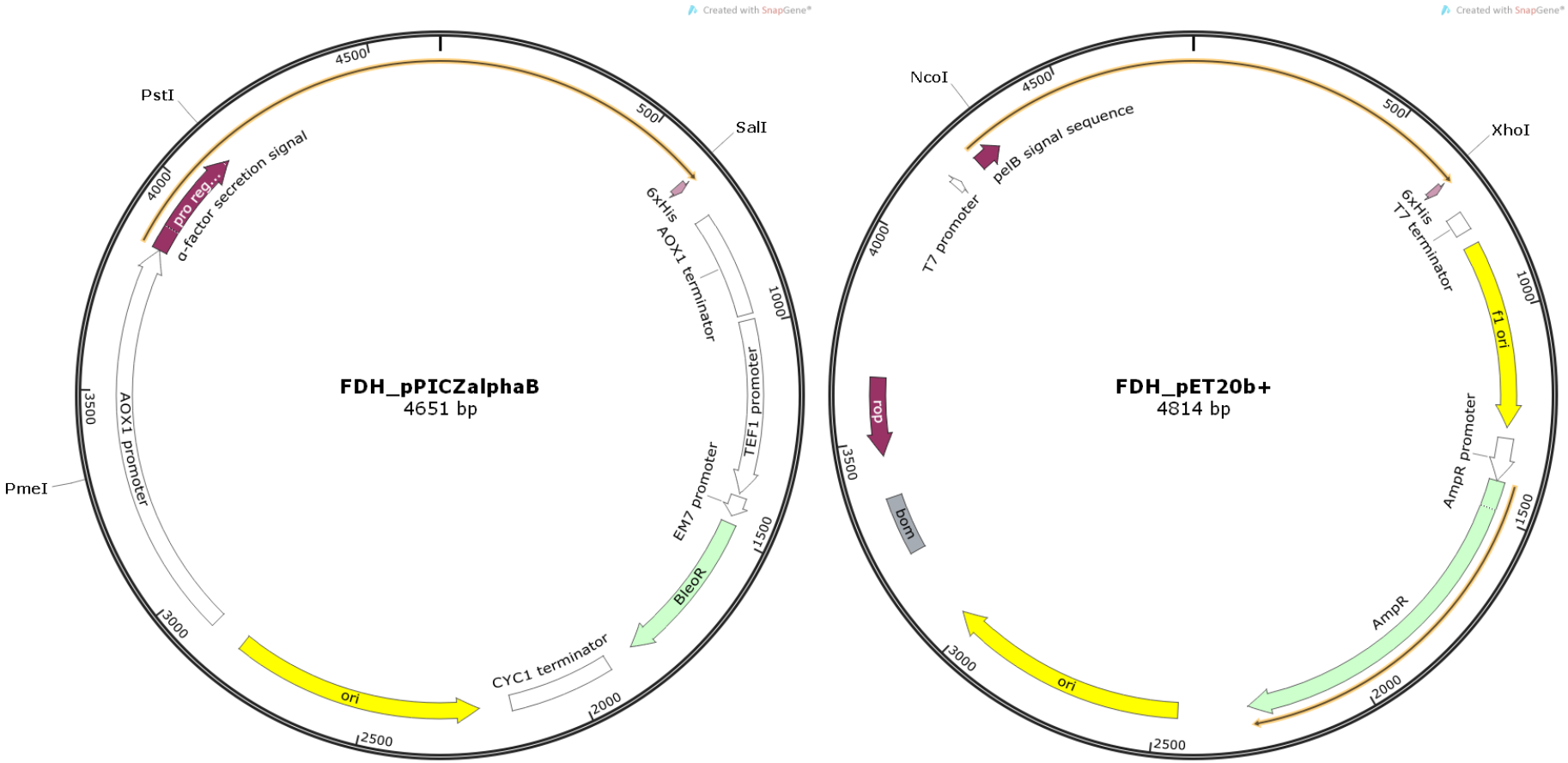

Figure A8: Expression vector encoding the formate dehydrogenase from *Burkholderia stabilis*. Left: Expression vector for the production and secretion in the recombinant host *Pichia pastoris*. Right: Expression vector for the production in the recombinant host *Escherichia coli* and translocation to the periplasm.

Table A4: Codon optimized sequences encoding formate dehydrogenase (*fdh*) from *Burkholderia stabilis* 15516 for heterologous expression in *Pichia pastoris* or *Escherichia coli*, respectively. UniProtKB entry FDH from *B. stabilis* 15516: B5A8W5.

Gene	Codon optimized sequence
fdh **(expression in *P. pastoris*)**	**CTGCAGCT** ATGGCTACTGTTTTGTGTGTTTTGTACCCAGATCCTGTTGATGGTTACCCACCTCATTACGTTAGAGATACTATCCCTGTT ATTACTAGATATGCTGATGGTCAAACTGCTCCAACTCCTGCTGGTCCACCTGGTTTTAGACCTGGTGAATTGGTTGGTTC TGTTTCTGGTGCTTTGGGTTTGAGAGGTTACTTGGAGGCTCATGGTCACACTTTGATTGTTACTTCTGATAAGGATGGTC CTGATTCTGAATTTGAGAGAAGATTGCCAGATGCTGATGTTGTTATTTCTCAACCTTTCTGGCCAGCTTATTTGACTGCT GAAAGAATTGCTAGAGCTCCTAAATTGAGATTGGCTTTGACTGCTGGTATTGGTTCTGATCATGTTGATTTGGATGCTGC TGCTAGAGCTCACATTACTGTTGCTGAAGTTACTGGTTCTAACTCTATTTCTGTTGCTGAGCATGTTGTTATGACTACTTT GGCTTTGGTTAGAAACTACTTGCCATCTCACGCTATTGCTCAACAAGGTGGTTGGAACATCGCTGATTGTGTTTCTAGAT CTTACGATGTTGAGGGTATGCATTTTGGTACTGTTGGTGCTGGTAGAATTGGTTTGGCTGTTTTGAGAAGATTGAAGCCT TTCGGTTTGCATTTGCACTACACTCAAAGACATAGATTGGATGCTGCTATTGAACAAGAGTTGGGTTTGACTTATCACG CTGATCCAGCTTCTTTGGCTGCTGCTGTTGATATCGTTAACTTGCAAATCCCTTTGTACCCATCTACTGAGCATTTGTTTG ATGCTGCTATGATTGCTAGAATGAAGAGAGGTGCTTATTTGATTAATACTGCTCGTGCTAAGTTGGTTGATAGAGATGC TGTTGTTAGAGCTGTTACTTCTGGTCACTTGGCTGGTTACGGTGGAGATGTTTGGTTTCCACAACCTGCTCCAGCTGATC ATCCTTGGAGAGCTATGCCATTCAACGGTATGACTCCTCACATTTCTGGTACTTCTTTGTCTGCTCAAGCTAGATATGCT GCTGGTACTTTGGAAATTTTGCAATGTTGGTTCGATGGTAGACCAATTAGAAATGAGTACTTGATTGTTGATGGTGGTA CTTTGGCTGGTACTGGTGCTCAATCTTATAGATTGACT **GTCGAC**
fdh **(expression in *E. coli*)**	**CCATG** **G**CGACCGTGCTGTGCGTTCTGTATCCGGACCCGGTGGATGGTTACCCGCCGCACTATGTGCGTGACACCATCCCGGTTA TTACCCGTTATGCGGATGGTCAAACCGCGCCGACCCCGGCGGGTCCGCCGGGCTTCCGTCCGGGCGAGCTGGTTGGTAG CGTGAGCGGTGCGCTGGGCCTGCGTGGTTACCTGGAAGCGCACGGCCACACCCTGATCGTGACCAGCGACAAGGATGG TCCGGACAGCGAGTTCGAACGTCGTCTGCCGGACGCGGATGTGGTTATCAGCCAACCGTTTTGGCCGGCGTATCTGACC GCGGAACGTATTGCGCGTGCGCCGAAACTGCGTCTGGCGCTGACCGCGGGTATTGGTAGCGATCATGTTGATCTGGATG CGGCGGCGCGTGCGCACATCACCGTGGCGGAAGTGACCGGTAGCAACAGCATTAGCGTGGCGGAACACGTGGTTATGA CCACCCTGGCGCTGGTTCGTAACTACCTGCCGAGCCATGCGATTGCGCAGCAAGGTGGCTGGAACATTGCGGACTGCGT GAGCCGTAGCTATGATGTTGAGGGCATGCACTTCGGTACCGTGGGTGCGGGTCGTATTGGCCTGGCGGTTCTGCGTCGT CTGAAGCCGTTTGGTCTGCACCTGCACTACACCCAACGTCACCGTCTGGATGCGGCGATCGAGCAAGAACTGGGCCTGA CCTATCATGCGGACCCGGCGAGCCTGGCGGCGGCGGTGGATATCGTTAACCTGCAGATTCCGCTGTACCCGAGCACCG AACACCTGTTCGATGCGGCGATGATCGCGCGTATGAAGCGTGGTGCGTATCTGATTAACACCGCGCGTGCGAAACTGGT GGACCGTGATGCGGTGGTGCGTGCGGTTACCAGCGGTCACCTGGCGGGTTATGGTGGCGACGTGTGGTTCCCGCAACC GGCGCCGGCGGATCACCCGTGGCGTGCGATGCCGTTTAACGGCATGACCCCGCACATTAGCGGTACCAGCCTGAGCGC GCAGGCGCGTTATGCGGCGGGTACCCTGGAGATCCTGCAATGCTGGTTTGACGGTCGTCCGATCCGTAACGAATACCTG ATTGTTGATGGTGGCACCCTGGCGGGTACCGGTGCGCAGAGCTATCGTCTGACC **CTCGAG**

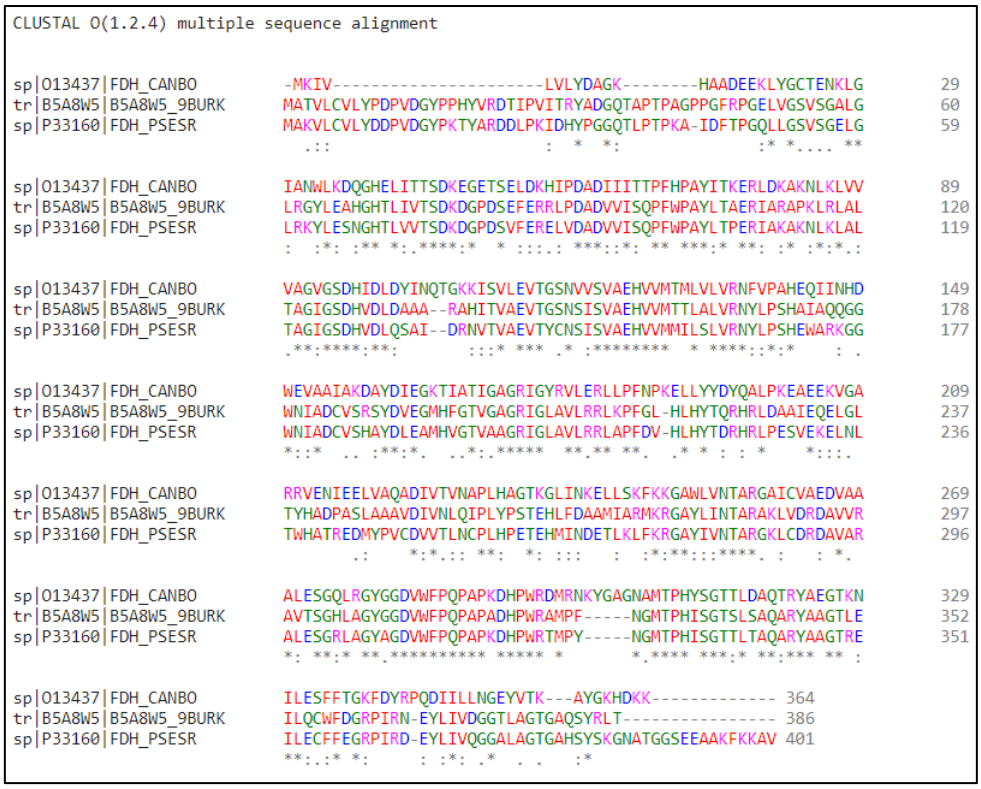

Figure A9: Alignment of FDH from *Burkholderia stabilis* to the extensively investigated FDH from *Candida boidinii* and *Pseudomonas sp. 101*.