

**Mining the Arabidopsis genome  
for cytochrome P450  
biocatalysts**

**Maria Magdalena Razalan**

**PhD**

**University of York**

**Biology**

**September 2016**

## Abstract

Cytochromes P450 (CYPs) constitute a wide group of NAD(P)H-dependent monooxygenases, found throughout all kingdoms of life. Among the most important functions of CYPs are the synthesis of bioactive compounds and the conversion of xenobiotics. These functions can be translated into biotechnological applications, such as the production of highly regio- and stereo-specific drug metabolites for the pharmaceutical industry, or to confer activity towards toxic compounds for agronomics and bioremediation purposes.

Plants possess a large number of CYP sequences, but most still remain uncharacterised, due to the difficulty in the isolation of the membrane-bound enzyme and in the reconstitution of an active and efficient redox system.

In this project, a fusion construct for the co-expression of the CYP with a suitable reductase was created. The construct consisted of a C-terminal Arabidopsis ATR2 reductase (codon-optimised for expression in *E. coli* and truncated of the N-terminal membrane anchor) connected through a poly-GlySer linker to the heme domain. An N-terminal Im9 peptide replaced the natural membrane-binding domain of the CYP. When CYP73A5 from Arabidopsis was cloned into the construct, it was able to convert almost 60 % of the substrate cinnamic acid to the hydroxylated derivative, in whole cell assays. This result demonstrated that this expression platform enables the expression of active redox self-sufficient P450 catalysts and it can be further utilised for the characterisation of orphan CYPs.

Following from gene expression studies and reports on the existence of oxidative derivatives of TNT, the potential involvement of CYP81D11 in the detoxification of TNT was explored with different *in planta* assays, employing transgenic Arabidopsis lines and tobacco leaf discs. The results obtained were contrasting and did not provide a clear picture on the role of CYP81D11. Further studies have to be carried out in the future, using CYP81D11-knockout lines, as well as the purified enzyme.

# List of Contents

<b>Abstract</b> .....	<b>2</b>
<b>List of contents</b> .....	<b>3</b>
<b>List of tables</b> .....	<b>8</b>
<b>List of figures</b> .....	<b>9</b>
<b>Acknowledgements</b> .....	<b>12</b>
<b>Declaration</b> .....	<b>13</b>
<b>Chapter 1: Introduction</b> .....	<b>14</b>
1.1 History of cytochrome P450 enzymes: the discovery.....	14
1.2 Structure, nomenclature and organisation of CYP enzymes .....	18
Structure.....	18
Nomenclature .....	22
Organisation .....	23
Cytochrome P450 redox partners .....	25
1.3 Catalytic cycle of cytochromes P450.....	31
1.4 Human Cytochromes P450.....	33
1.5 Bacterial Cytochromes P450 .....	37
1.6 Plant Cytochromes P450 .....	42
1.7 Fungal Cytochromes P450.....	48
1.8 Challenges of using CYPs for biotechnological applications .....	49
Aim of the project .....	50
<b>Chapter 2: General Materials and Methods</b> .....	<b>51</b>
2.1 Chemical reagents.....	51
2.2 Organisms.....	51
2.3 Media .....	52

Media for bacterial growth .....	52
Media for plant growth and assays .....	53
2.4 Molecular Biology techniques.....	54
2.4.1 Polymerase Chain Reaction (PCR) & colony PCR.....	54
2.4.2 DNA restriction digestion .....	54
2.4.3 Dephosphorylation of linearised vector .....	54
2.4.4 Agarose gel electrophoresis .....	54
2.4.5 DNA isolation.....	55
2.4.6 Cloning: ligation, InFusion .....	56
2.4.7 Preparation of competent cells.....	56
2.4.8 DNA transformation .....	57
2.4.9. DNA sequencing .....	57
2.5 Biochemistry techniques.....	58
2.5.1 Protein heterologous expression .....	58
2.5.2 Cell harvest & lysis.....	58
2.5.3 Spectrophotometric evaluation of protein expression .....	59
2.5.4 SDS-PAGE verification.....	59
2.5.5 Western blot verification.....	60
2.5.6 Protein purification .....	60
2.5.7 Protein quantification .....	62
2.5.8 Activity assay .....	62
<b>Chapter 3: Soluble expression of Cytochromes P450.....</b>	<b>63</b>
3.1 Introduction: Bottlenecks in the heterologous expression of plant CYPs .....	63
3.2 Materials and methods used for the soluble expression of plant CYPs .....	68
3.2.1 Sequence analysis, PCR and cloning .....	68
3.2.2 Expression trials in E. coli: Rosetta 2, BL21(DE3), Arctic express .....	72

3.2.3 Expression trials in yeast ( <i>Saccharomyces cerevisiae</i> WAT11 modified strain) .....	72
Media and buffers for yeast growth and harvest.....	73
3.2.4 Purification .....	74
3.2.5 Protein characterisation: UV-Vis spectrophotometry and activity .....	74
3.3 Results .....	75
3.3.1 Sequence analysis and cloning.....	75
3.3.2 Expression trials .....	82
3.3.3 Expression in yeast.....	85
3.3.4 Purification of the S-CYP73A5tr construct .....	86
3.3.5 Activity assay .....	90
3.4 Discussion.....	90
Gene expression in bacteria.....	91
Gene expression in yeast.....	92
<b>Chapter 4: Expression screening and characterisation of P450- reductase fusion proteins</b>	<b>94</b>
4.1 Introduction: plant P450-reductase fusions .....	94
4.2 Materials and Methods used for the expression screening of the plant fusions .....	99
4.2.1 Cloning steps .....	99
4.2.2 First expression trials.....	101
4.2.3 High-throughput automated expression screening .....	102
4.2.4 Purification of the fused construct .....	105
4.2.5 Activity assay of the fused construct CYP73A5-ATR2.....	105
4.3 Results .....	106
4.3.1 Creation of the pAthIA2 fusion platform .....	106
4.3.2 First expression trial of the pAthIA2 (S-CYP73A5tr-ATR2tr) fusion in <i>E. coli</i> Arctic express .....	109

4.3.3 Re-cloning and automated protein expression screening .....	111
4.3.4 Purification of the Im9-CYP73A5 fusion construct.....	113
4.3.5 Activity assay of the fused Im9-CYP73A5tr-ATR2tr construct.....	114
4.4 Discussion.....	116
<b>Chapter 5: A role for CYP81D11 in the detoxification of the explosive TNT .....</b>	<b>119</b>
5.1 Introduction on the role of Cytochromes P450 in plant detoxification processes.....	119
5.2 Materials and Methods.....	127
5.2.1. Plants and growth media .....	127
Plants.....	127
Growth media .....	127
5.2.2. Sterilisation of Arabidopsis seeds .....	128
5.2.3 Verification of gene expression in the Arabidopsis CYP81D11-modified lines via qPCR.....	128
5.2.4. Hydroponic cultures of Arabidopsis transgenic plants & HPLC analysis of TNT uptake .....	129
5.2.5. Root length on-plate comparison for differential TNT-resistance.....	130
5.2.6. Transient expression of CYP81D11 in Nicotiana benthamiana .....	130
5.3 Results .....	134
5.3.1 Verification of CYP81D11 expression in Arabidopsis .....	134
5.3.2 TNT uptake by hydroponic Arabidopsis cultures.....	135
5.3.3 Comparison of CYP81D11-modified plants' root length.....	135
5.3.4 Transient expression of CYP81D11 in tobacco leaves .....	138
5.4 Discussion .....	143
<b>Chapter 6: Final discussion .....</b>	<b>149</b>
Future perspectives.....	153
Fusion platform .....	153

Elucidation of the role of CYP81D11 in TNT detoxification.....	154
<b>Abbreviations.....</b>	<b>155</b>
<b>References .....</b>	<b>157</b>

## List of Tables

<b>Table 1.1:</b> Number of annotated CYP sequences .....	14
<b>Table 1.2:</b> Human CYPs subdivided by major substrate specificity .....	34
<b>Table 1.3:</b> Bacterial CYP activities .....	37
<b>Table 2.1:</b> Specifications of the bacterial strains used in the present work.....	52
<b>Table 3.1:</b> Gene ID and locus tags of the CYPs used in the chapter .....	68
<b>Table 3.2:</b> Primers for the PCR amplification of the CYP targets.....	70
<b>Table 4.1:</b> Primers for the PCR amplification of ATR2 and S-CYP73A5tr.....	100
<b>Table 4.2:</b> List of primers for the cloning into the new expression vectors.....	103
<b>Table 4.3:</b> Structure of the pETFPP_1-5 expression vectors.....	104
<b>Table 5.1:</b> List of the qPCR primers.....	129
<b>Table 5.2:</b> Primers used for the cloning steps in the pK2GW7 plasmid .....	131

## List of Figures

<b>Figure 1.1:</b> CYP absorption spectra from Omura and Sato's 1962 publication .....	16
<b>Figure 1.2:</b> Number of papers on CYP enzymes published every year .....	18
<b>Figure 1.3:</b> Alignment of different cytochrome P450 enzymes.....	20
<b>Figure 1.4:</b> Topology and structure of CYP enzymes .....	21
<b>Figure 1.5:</b> CYP/CYP reductase configurations .....	24
<b>Figure 1.6:</b> CPR Structure and topology.....	26
<b>Figure 1.7:</b> Interaction between the domains in the CPR system .....	27
<b>Figure 1.8:</b> Human cytochrome P450 reductase. ....	28
<b>Figure 1.9:</b> Electron transport chain in <i>P. putida</i> P450cam .....	29
<b>Figure 1.10:</b> Electron transport chain in the AdR/AdX-CYP system.....	30
<b>Figure 1.11:</b> P450 catalytic cycle .....	32
<b>Figure 1.12:</b> Drug-metabolising P450 enzymes .....	34
<b>Figure 1.13:</b> Structure of <i>B. megaterium</i> P450 BM-3 .....	38
<b>Figure 1.14:</b> Illustration of the LICRED drop-in platform.....	40
<b>Figure 1.15:</b> CYP71AV1 in the synthesis of artemisinin .....	44
<b>Figure 1.16:</b> CYP725A1 in the taxol biosynthetic pathway.....	45
<b>Figure 1.17:</b> The <i>S. bicolor</i> dhurrin biosynthetic pathway .....	46
<b>Figure 3.1:</b> Reaction catalysed by CYP73A5.....	65
<b>Figure 3.2:</b> Expression pattern of AtCYP81D11 at different developmental stages..	66
<b>Figure 3.3:</b> Expression pattern of AtCYP81D11 in different tissues .....	67
<b>Figure 3.4:</b> Plasmid maps for pET28a and for the synthetic pMA-T plasmid .....	71
<b>Figure 3.5A:</b> Transmembrane domain predictions for AtCYP73A5. ....	76
<b>Figure 3.5B:</b> Transmembrane domain predictions for AtCYP81D11. ....	77
<b>Figure 3.5C:</b> Transmembrane domain predictions for ZmCYP81A9. ....	78
<b>Figure 3.6:</b> PCR amplification of CYP73A5tr and CYP81D11tr .....	79
<b>Figure 3.7:</b> Double digestion of the pMA-T vector .....	80
<b>Figure 3.8:</b> Colony PCR verification.....	81
<b>Figure 3.9:</b> Second set of PCR amplifications.....	81
<b>Figure 3.10:</b> Double digestion of pET28a .....	82
<b>Figure 3.11:</b> SDS-PAGE analysis of the expression of S-CYP73A5tr .....	83

<b>Figure 3.12:</b> SDS-PAGE analysis of the expression of S-CYP73A5tr .....	84
<b>Figure 3.13:</b> Western blot analysis of the expression trials of S-CYP73A5tr .....	84
<b>Figure 3.14:</b> Western blot analysis of the expression of the S-CYPtr constructs .....	85
<b>Figure 3.15:</b> CO-binding spectra of the yeast microsomes.....	86
<b>Figure 3.16:</b> SDS-PAGE of the samples from the purification of S-CYP73A5tr .....	87
<b>Figure 3.17:</b> UV-Vis absorbance spectrum of the purified S-CYP73A5tr. ....	87
<b>Figure 3.18:</b> SDS-PAGE of the samples from the 2 <sup>nd</sup> purification of S-CYP73A5tr. ...	88
<b>Figure 3.19:</b> Sample colouration .....	88
<b>Figure 3.20:</b> Spectrophotometric characterisation of the purified S-CYP73A5.....	89
<b>Figure 4.1:</b> Structural configuration of <i>B. megaterium</i> P450-BM3.....	94
<b>Figure 4.2:</b> Map of the ACryLAM/LIC fusion platform .....	97
<b>Figure 4.3:</b> Map of the fusion platform. ....	101
<b>Figure 4.4:</b> Overview of the automated process for protein expression screening	102
<b>Figure 4.5:</b> Map of the pET-YSBL3C vector .....	103
<b>Figure 4.6:</b> Transmembrane domain predictions for AtATR2.....	107
<b>Figure 4.7:</b> Agarose gel analysis of the PCR amplification of ATR2tr.....	108
<b>Figure 4.8:</b> Colony PCR verification of ATR2 cloning .....	108
<b>Figure 4.9:</b> Amplification of the S-CYP73A5tr insert.....	109
<b>Figure 4.10:</b> SDS-PAGE analysis of the expression trial of pAtHIA2.....	110
<b>Figure 4.11:</b> Western blot analysis of the expression of ATR2tr. ....	110
<b>Figure 4.12:</b> Dot blot/Western blot analysis from the expression trial.....	111
<b>Figure 4.13:</b> SDS-PAGE analysis of the first expression screening.....	112
<b>Figure 4.14:</b> Overlap of the absorbance spectra .....	112
<b>Figure 4.15:</b> SDS-PAGE and western blot analysis of the Im9-fusion .....	113
<b>Figure 4.16:</b> SDS-PAGE of the fractions from the purification of the Im9 fusion ...	113
<b>Figure 4.17:</b> HPLC analysis of the conversion of cinnamic acid.....	115
<b>Figure 5.1:</b> Conversion of xenobiotic compounds in plant cells .....	120
<b>Figure 5.2:</b> Chemical structures of TNT and RDX.....	122
<b>Figure 5.3:</b> Structure of the XplA-XplB system.....	122
<b>Figure 5.4:</b> RDX degradation by XplA (CYP177A1). ....	123
<b>Figure 5.5:</b> Structure of the XplA heme domain .....	124
<b>Figure 5.6:</b> Steps in the biotransformation of TNT. ....	125

<b>Figure 5.7:</b> Map of the pK2GW7 binary vector .....	131
<b>Figure 5.8:</b> qPCR results from the TNT-treated Arabidopsis plants.....	134
<b>Figure 5.9:</b> TNT concentration in the plants' growth medium .....	135
<b>Figure 5.10 a-e:</b> Pictures of 20-day-old Arabidopsis plants.....	137
<b>Figure 5.11:</b> Root lengths of 20-day old Arabidopsis seedlings.....	137
<b>Figure 5.12:</b> Western blot analysis of CYP81D11 expression in tobacco .....	138
<b>Figure 5.13:</b> TNT uptake in the tobacco leaf discs incubations .....	139
<b>Figure 5.14:</b> TNT derivatives in the tobacco leaf discs incubations.....	139
<b>Figure 5.15:</b> Pictures of the tobacco leaf discs post TNT-incubation .....	140
<b>Figure 5.16:</b> Chromatographic profiles of the tobacco leaf samples .....	140
<b>Figure 5.17:</b> Comparison of the peak areas of the tobacco tissue extracts. ....	141
<b>Figure 5.18:</b> MS-ESI spectra of the metabolites extracted from the tissues.....	142
<b>Figure 5.19:</b> TNT derivatives from Arabidopsis tissues.....	143

# Acknowledgements

These last four years as a PhD student have been a challenging yet pleasant adventure for me.

I can't be thankful enough to Prof. Neil Bruce for the guidance and support throughout these years. A very special thank you goes also to Dr. Liz Rylott and Margaret, for their precious help (and patience!) in the lab and in the office, respectively.

I would like to thank Dr. James Moir and Prof. Gideon Grogan for the constructive discussions at the periodic TAP meetings, Prof. Johnathan Napier (Rothamsted Research) for providing the Arabidopsis transgenic lines, Luisa Elias (CNAP, McQM group) for the tobacco seeds, Dr. Jared Cartwright and Mick Miller (BioScience Technology Facility) for the automated expression screening platform and Dr. Swen Langer (CNAP Metabolite Profiling Unit) for the MS analyses.

I really enjoyed my involvement in the P4fifty network, it was such a great international cluster of researchers with whom I had the privilege to work for three years. In particular, I'd like to thank Prof. Birger Møller and Prof. Danièle Werck for hosting me at their respective laboratories.

A huge '**THANK YOU**' goes

- to the whole CNAP-M2 group, to every single colleague that inspired me, encouraged me... and fed me with delicious cakes!
- to the greenSTEMS bunch, especially Giulia P., Jen C., Ruth, Jennie D. and Erin, for making this PhD an enjoyable journey of discovery;
- to Claudia-san, my travel buddy, not only in the PhD adventure, but also a great fellow backpacker in many trips;
- to Karla, Maria, Alberto, Zak and Laz, for the comforting and supporting words throughout the final writing period;
- to Sara and Paolo, for the constant love and support, even across continents and oceans;
- to Kyriakos: a knowledgeable colleague and lovely partner, the most generous and patient person on Earth! I don't think I would have been able to cope with all the difficulties in this path without his helping hand and shoulders to lean on;
- to my amazing mom and dad: I will never be able to thank them enough for the mountain of sacrifices they've done for me in the last thirty years... they are my heroes!

## **Declaration**

I declare that I am the sole author of this work and that it is original except where indicated by reference in the text. No part of this work has been submitted for any other degree to any other institution.

# Chapter 1: Introduction

## 1.1 History of cytochrome P450 enzymes: the discovery

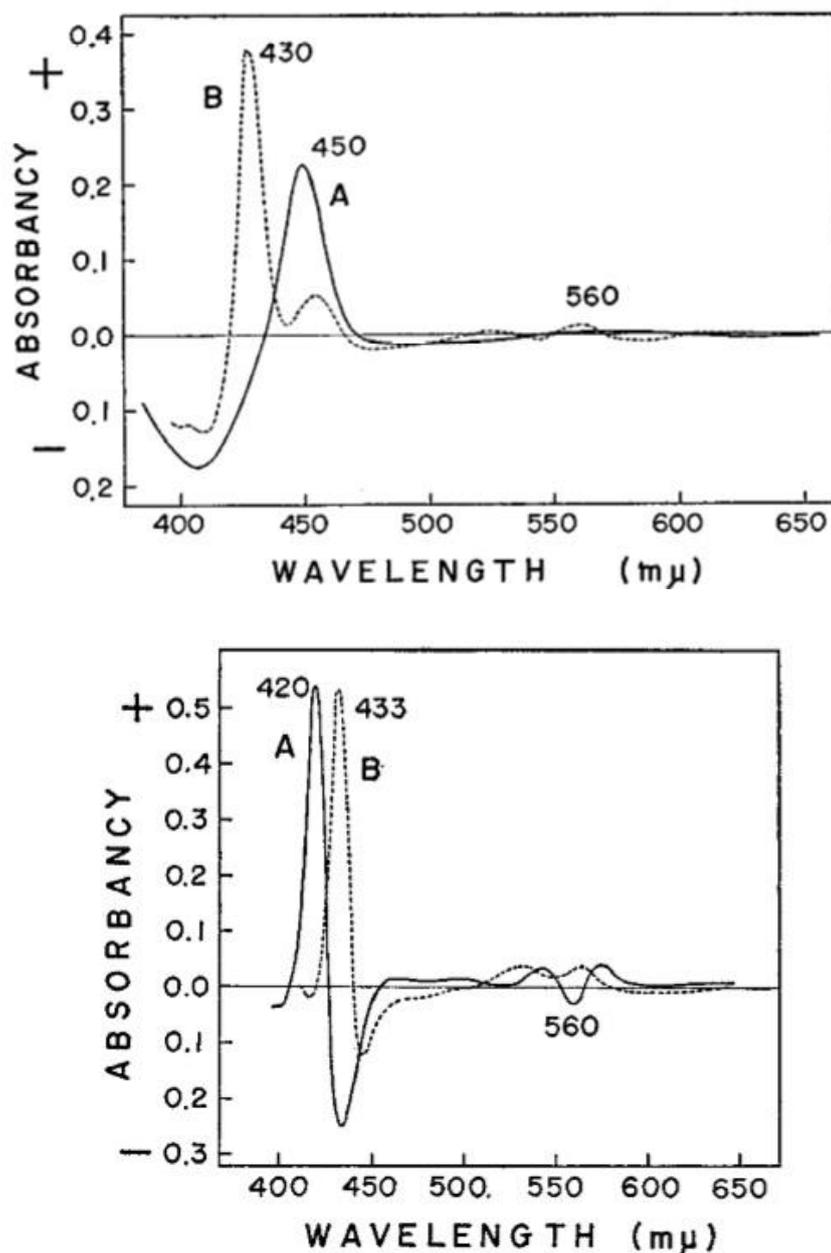
Cytochrome P450 (CYP) enzymes form a superfamily of heme-thiolate monooxygenases distributed ubiquitously across all kingdoms of life, with few exceptions. According to the latest count, published by Prof. David Nelson, University of Tennessee, in April 2016 (<http://drnelson.uthsc.edu/P450.statsfile.html>), more than 35,000 P450 sequences have been annotated, with the largest group presented by plants (> 13,000), followed by mammals (>10,000) and fungi (>7,800), as reported in table 1.1.

**Table 1.1:** Number of CYP sequences annotated as of April 2016, according to Prof. Nelson's latest updated count (<http://drnelson.uthsc.edu/P450.statsfile.html>, 09/09/2016), on the Cytochrome P450 homepage.

Organism	Number of P450 sequences
Plants	13,978
Mammalian	10,477
Fungi	7,873
Bacteria	2,156
Protozoa	602
Archaea	52
Viruses	28
<b>TOTAL</b>	<b>35,166</b>

The fifty-year old history of cytochrome P450 enzymes stemmed from the identification and characterisation studies on mammalian microsomal cytochrome b<sub>5</sub> (cyt<sub>b5</sub>) (1-3). Klingenberg, from the University of Pennsylvania, confirmed, a few years later, the preliminary observations made by Dr. G. R. Williams on rat liver microsomes. A microsomal compound, upon reduction with NADPH or dithionite and subsequent addition of carbon monoxide (CO), displayed a peculiar absorbance peak at 450 nm, different from what observed for cyt<sub>b5</sub> (4). In the same period, Garfinkel (a colleague of Klingenberg) conducted studies on liver microsomes from pigs,

instead, with a focus on electron transport chains. In the samples examined, a pigment could be separated using centrifugation from the well-known  $\text{cyt}_{b5}$ . This pigment could be reduced with dithionite and, by binding CO, it gave a characteristic absorbance peak at 450 nm (5). In 1962, Omura and Sato, from Osaka University, published the first paper focused on this CO-binding pigment. In this work, rabbit liver microsomal suspensions were employed, unveiling new characteristics of the compound, which was named 'P-450'. The dithionite-reduced P-450 could react not only with CO, but also with ethylisocyanide and nitric oxide; these reactions had been already previously described for other hemoproteins. These observations, in addition to the overabundance of recorded total heme (as compared to the amount of recorded  $\text{cyt}_{b5}$ ) in the microsomal sample, led to the speculation that the P-450 compound was in fact a hemoprotein. Omura and Sato reported also that the P-450 behaved in different ways in the presence or absence of oxygen. In anaerobic conditions, the compound could be reduced by NAD(P)H to the same extent as with dithionite. However, when oxygen was present, the P-450 pigment could be reduced effectively only with dithionite and was highly unstable thereafter, due to the peroxidation of the microsomal lipids and, subsequently, the degradation of the heme. In contrast to  $\text{cyt}_{b5}$ , cysteine and ascorbate did not reduce the P-450. Another peculiar property was observed when the microsomes were solubilised with snake venom (which contained phospholipase A) or deoxycholate, in anaerobic conditions. The release of the compound from the microsomal phospholipids induced a hypsochromic shift in the spectrum, leading to the disappearance of the 450 nm peak and the formation of a new absorption maximum, at 420 nm (figure 1.1) (6). The newly-formed compound was therefore named 'P-420'. When a difference spectrum of the dithionite-reduced minus the oxidised form of P-420 was recorded, the trace resembled that of other hemochromogens, with peaks at 427 nm, 530 nm and 560 nm, further proving the hemoprotein nature of the compound (7).



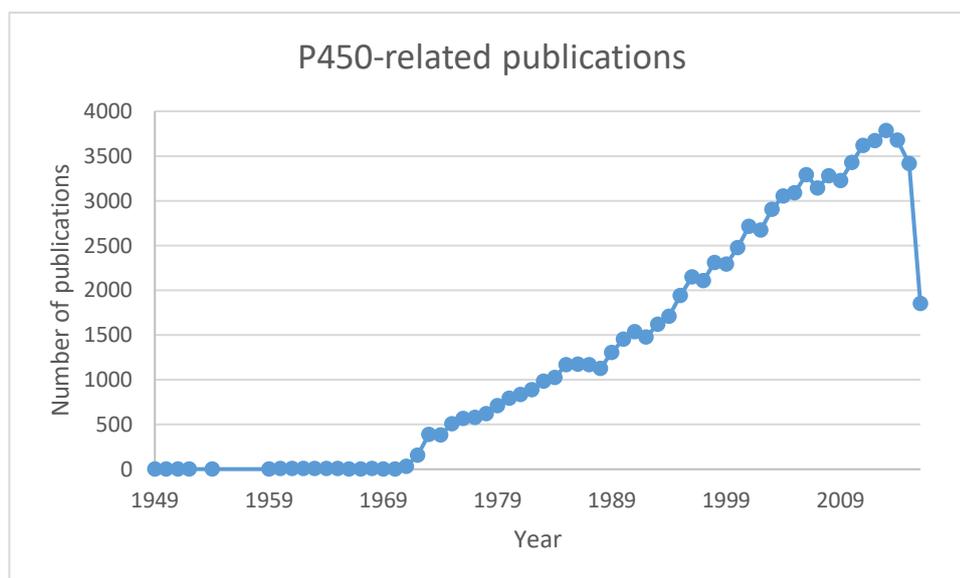
**Figure 1.1:** Original absorption spectra from Omura and Sato's 1962 publication (6). Top panel: spectra of the microsomal suspensions, reduced with sodium dithionite and treated with carbon monoxide for 20 s (curve A), or treated with 100  $\mu\text{M}$  ethylisocyanide (curve B). Bottom panel: spectra of the microsomal suspensions, digested with 1% snake venom and treated with carbon monoxide (curve A) or with ethylisocyanide (curve B).

Subsequently, Omura and Sato calculated the molar extinction coefficient of the P-450 compound ( $91 \text{ mM}^{-1} \text{ cm}^{-1}$ ) and purified its 'solubilised' counterpart (P-420), with a yield of 6-7  $\mu\text{moles}$  of protoheme/mg of total protein. The multistep purification process was preceded by two fractional solubilisation steps: first, the microsomes were digested with a lipase (pancreatic steapsin) and then solubilised

either with snake venom or deoxycholate (8). Nevertheless, it was observed that the P-420 compound could also form during the enzymatic digestion of the microsomes, meaning that the transition P-450 to P-420 was not exclusively due to microsome solubilisation (9).

Shortly thereafter, Cooper *et al.*, examining some earlier studies by Ryan and Engle focused on steroid synthesis in bovine adrenocortical microsomal fractions (10), and finally assigned a function to the P-450 compound. In fact, Cooper and co-workers demonstrated that the specific hydroxylating microsomal fraction contained a P-450 compound. Additionally, they showed that the catalysis was inhibited by CO and dependent not only on the presence of NADPH, but also of molecular oxygen. This observation proved valid also for other types of reactions occurring in rat microsomes, such as codeine demethylation and acetanilide hydroxylation, leading to the classification of the P-450 pigment from mammalian microsomes as an oxygenase. Furthermore, the same authors, observing that the incubation of the bovine C21 steroid hydroxylase P-450 with sulfhydryl inhibitors led to an irreversible spectral transition to P-420, concluded that the P-420 compound was an inactivated version of the P-450 enzyme (11). Gunsalus and co-workers were among the first researchers to focus on P450 reactions in bacteria, working with different strains of *Pseudomonas putida*. In this organism the first multi domain P450 system, named P450cam, was discovered, and characterised as responsible for the hydroxylation of camphor (12, 13). Other pioneering studies in the area of microbial CYP activities were carried out by Appleby, who described in 1967 the involvement of P-450 enzymes in nitrogen reduction by rhizobium bacteria (14).

Since then, thousands of P450-related papers have been published every year, as recorded by the National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/>, figure 1.2).



**Figure 1.2:** Number of papers on CYP enzymes published every year in the last 60 years, as recorded by the NCBI database.

## 1.2 Structure, nomenclature and organisation of CYP enzymes

### Structure

Cytochrome P450 enzymes are highly divergent at the primary sequence level with, in some cases, a sequence identity lower than 20 %. Only a few domains, in fact, are conserved throughout the species, mainly located around the active site, which is characterised by a porphyrinic heme ring, with a central iron atom coordinated with four nitrogen atoms, a cysteine (the only residue conserved among all P450s) and a water molecule. The most important consensus sequences are Ala/Gly-Gly-X-Ala/Glu-Thr-Thr/Ser (responsible for oxygen binding and activation) in the core of the I helix, which runs over the distal face of the heme, Phe-XX-Gly-X-Arg-X-Cys-X-Gly (the Cys is the heme iron ligand) and Glu-X-X-Arg, both situated at the proximal side of the heme. At the N-terminus of plant and animal P450 sequences a proline-rich segment, X-Pro-Gly-Pro, can often be found, followed by Pro-X-X-Gly (see alignments in figure 1.3) (15-17).

## Chapter 1: Introduction

RrXp1A	MTDVTVLFGTETGNAEMVADDIASALGEFDIEATVVGMEFD--VADLAASG---TVVLV	55
PpP450cam	-----	0
HsCYP19A1	-----MVLEMLNP--IHYNITSIVPEAMPA-ATMPVLLLT---GLFLLV	38
BmP450-BM3	-----	0
HsCYP3A4	-----MAL---IPDLAMETWLLLAVSLVLLLYLYG	26
HsCYP2D6	-----MGLEALVPLAVIVAIIFLLLVLDL	22
AtCYP73A5	-----MDLLLLLEKSLIAVFAVILATV	22
HtCYP76B1	-----MDFLIIIVSTLLLS-----YILI	17
AaCYP71AV1	-----MKSIL-KAMALS-LTTSIALATI-LLFVYK	27
RrXp1A	TSTYG---EGELPA-TTQPFDFAMKAAEPDLT-GLRFGAFLGDSTYDTYNNAIDILVGA	110
PpP450cam	-----	0
HsCYP19A1	WNYEG---TSSIPGGPGYCMGIGPLISHGRFLWMGIGSACN---YYNRVYGEFMRVWISG	91
BmP450-BM3	-----MTIKEMPQKTFGELKNLPLLNTDKP---VQALM---KIADELGEIFKFEAPG	47
HsCYP3A4	THSHGLFKKLGIPGPTPLPFLGNILSYHKGFC---MFD-M---ECHKKYGVWGFYDGGQ	78
HsCYP2D6	MHRRQRWAARYPPGPLPLPGLGNLLHVDQNT---PYCFD---QLRRRFGDVFSLQLAW	75
AtCYP73A5	ISKLRGKLLKLPPIPIPIFGNWLQVGDLDL---HRNLV---DYAKKFGDLFLLRMGQ	75
HtCYP76B1	WVLGVGPKNLPPIPIPIIIGNLH-LLGALP---HQSLA---KLAKIHGPIMSLQLGQ	69
AaCYP71AV1	FATRSKSTKSLPEPWRLPIIGHMHHLIGTTP---HRGVR---DLARKYGSMLHLQLGE	80
RrXp1A	VT-----DAGAT-QVGATGRHDAA-SFQPA--DGPVAEWAKQFAEALSDRTRRGHEMTA	161
PpP450cam	-----MTTETIQSNANLAPLPPHVPEHLV-F-----DFDMYNPSN	34
HsCYP19A1	EETLIISKSSSMFHMKHNHYSRFG-S-----KLGQL-----CIGMHEKGI	132
BmP450-BM3	RVTRYLSSQRLIKEACDESFRDKNLS-QAL---KFVR-----DFAGDGLF-	88
HsCYP3A4	QPVLAITDPDMIKTVLVKECYVFTNRRPF---GPV-----GFM---KSA	117
HsCYP2D6	TPVVVLNGLAAVREALVTHGEDTA-D-RPP---VPITQI-L-----GFGPRSQGV	119
AtCYP73A5	RNLVVVSSPDLTKEVLLTQGEVFG-S-RTR---NVVFDI-----FTGKGQDM	117
HtCYP76B1	ITTLVISSATAAEVLLKKQDLAFS-T-RNV---PDAVRA-----YNHERHSI	111
AaCYP71AV1	VPTIVVSSPKWAKEILTTYDITFA-N-RPE---TLTGEI-----VLYHNTDV	122
RrXp1A	ASIDRELVPSWDP-EFRNNPYPWYRRLQQDHPVHKLEDGTYLVSRVADVSHFAKLPIMSV	220
PpP450cam	-LSAGVQEAWAVLQESNVPDLVWTRCNG---GHWIATRQGLIREAYEDYRHSSECP---	87
HsCYP19A1	-IFNNNPELWKTTRPFPMKALSGPGLVR-----MVTCAES-----LKT---	170
BmP450-BM3	-TSWTHEKNWKAHNILLPSFSQQAMKG-YHAMMV-DIAVQ-----LVQ---	129
HsCYP3A4	-ISIAEDEEWKRLRSLLSPTFTSGKLIKE-MVPIIA-QYGDV-----LVR---	158
HsCYP2D6	-FLARYGPAWREQRRFSVSTLRNLGLGKKSLEQWTEEAAC-----LCA---	162
AtCYP73A5	-VFTVYGEHWKRRIMTVPFNTKVVQQNREGWFEAAS-----VVE---	159
HtCYP76B1	-SFLHVCTEWRTLRRIVSNIFSNSSLKQHLRSKKEVE-----LIA---	153
AaCYP71AV1	-VLAPYGEYWRQLRKICTLELLSVKVKVSFQSLREEECWN-----LVQ---	164
	*	
RrXp1A	EPGWADAGPWAVASDTALGS-----DPPHHTVLRRTQNKWFTPKLVDGWVRTTRELVDGL	275
PpP450cam	--FI-----PREAGEAYDFIPTSMPPPEQRQFRALANQVVGMPVVDKLENRIQELACSL	139
HsCYP19A1	--HL-----DRLEEVTNESGYVD-----VLTLLRRVMLD-----TSNTLFLRIP	207
BmP450-BM3	--KW-----ERLN-ADEHIEVPE--DMTRLTLDITGLCGFNRYRFSFYRDQPHPFITSM	178
HsCYP3A4	--NL-----RREAETGKPVTLKD--VFGAYSMDVITSTSFQVNIIDS-LNNPQDPFVENT	207
HsCYP2D6	--AF-----ANHSGRPFPRPN-----GLLDKAVSNVIAASLTCGRRFEYDDPRFLRL-	205
AtCYP73A5	--DV-----KKNPDSATKGIIVLR--KRL---QLMMYNNMFRIMFDRRFESEDDPLFLRL	206
HtCYP76B1	--YC-----RKAALSNENVHIGR--AAFRTSLNLLSNTIFSKDLTDPYEDSGKEFREV-	202
AaCYP71AV1	--EI-----KASGSGRPVNLSE--NVFKLIATILSRAAFGKGIKD----QKELTEI-	207
	;	
RrXp1A	LDGVEA-GQVIEA-----RRDL-AVVPHTVMARVLQLPEDDADAVMEAMFEAMLMQSA	327
PpP450cam	IESLRPQGCNFT-----EDYAEFPPIRIFMLL-AGLPEEDIPHLK---YLTD--QMT	186
HsCYP19A1	LD--ESAIIVVKIQGYFDWAQALLIKPDIFFKISW-----LYK	242
BmP450-BM3	VRALDEAMN-KL-----QRANPDDP-----AYD	200
HsCYP3A4	-----KK-LLRFDLDPFFLSITVFPFLIPILE-----VLN	237
HsCYP2D6	LD-LAQEGL-KEESGFL---REVLNAVVPVLLHIPA-----LAG	238
AtCYP73A5	KA-LNGERS-RLAQSFSE---YNYGDFIPIILRPFLR-----GYL	239
HtCYP76B1	-----ITN-IMVDSAK---TNLVDFVPLKIDP-----QGI	230
AaCYP71AV1	-----VKE-ILRQTGG---FDVADIFPSKFLHH-----LSG	235

**Figure 1.3:** Alignment of different cytochrome P450 enzymes (continues on the following page)

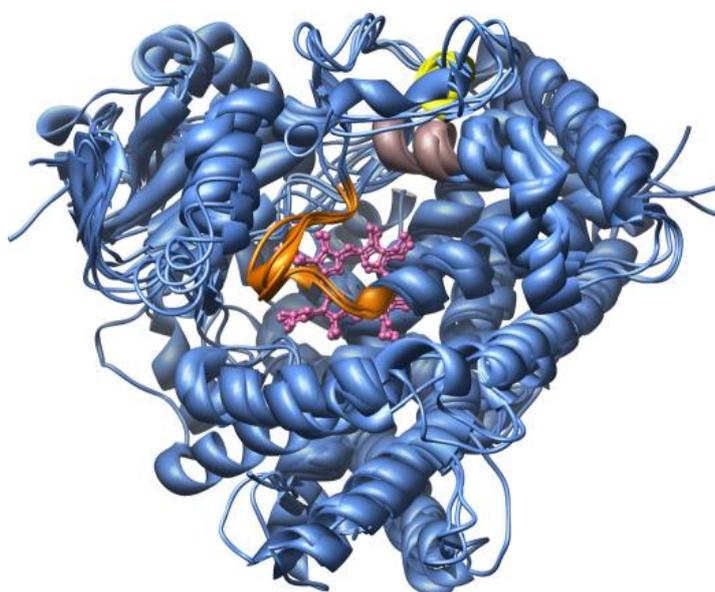
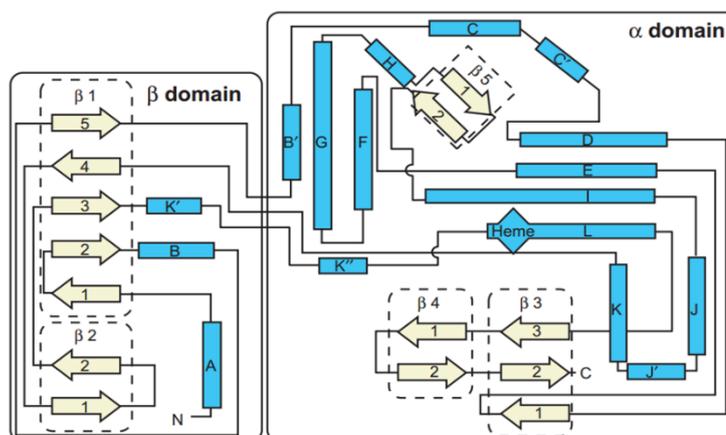
## Chapter 1: Introduction

RrXp1A	EPADG-DV-DRAAVAFGYLSARVAEM-LEDKRVN-----PGDGLADSLLD-----	369
PpP450cam	RPDGSMTFAEAKEALYDYL----IPI-IEQRRQK-----PGTDAISIVANGQ----	228
HsCYP19A1	K-----YEKSVKDLKDAI----EVL-IAEKRRRISTEE-KLEECCMFATELI-----L	284
BmP450-BM3	E--NKRQFQEDIKVMNDLV----DKI-----IADRK-ASGEQSDDLLTHMLNGK----D	243
HsCYP3A4	I---CV-FPRE---VTNFL----RKS---VKRMKESRLEDTQKHRVDFLQLMIDSQNSKE	283
HsCYP2D6	K---VLRFAQK---FLTQL----DEL-LTEHRMTWDPAPPRDLTEAFLAEMEKAKG---	284
AtCYP73A5	K--ICQDVKDR---RIALF----KKYFVDERKQIASSKPTGSEGLKCAIDHIL--EA---	285
HtCYP76B1	KRGMARHF-SK---VLGIF----DQL-IEE-RMR-----TGRFEQGDVLDVCLKMMQ---	272
AaCYP71AV1	KRARLTSLRKK---IDNLI----DNL-VAEHTVN-----TSSKTNETLLDVLRLKLD---	279
	.	
RrXp1A	AARAGEITESEAIATILVFYAVGHMAI GYLIASGIELFARRPEVFTAFRNDESAR-----	424
PpP450cam	-VNGRPITSDEAKRMCGLLLVGGLDVTNFLSFSMEFLAKSPEHRQELIERPERIPAA--	285
HsCYP19A1	AEKRGDLTRENVNQCILEMLIAAPDTMSVSLFFMFLIAKHPNVEEAIKEIQTVIGER-	343
BmP450-BM3	PETGEPLDDENIRYQIITFLIAGHETTSGLLSFALYFLVKNPHVLQKAAEEAARVLVDP-	302
HsCYP3A4	TESHKALSDLELVAQSIIIFIFAGYETTSSVLSFIMYELATHPDVQQKLEEEIDAVLPNKA	343
HsCYP2D6	-NPESFNDENLRIVVADLFSAGMVTSTTLAWGLLLMILHPDVQRRVQVEIDDVIGQVR	343
AtCYP73A5	-EQKGEINEDNVLYIVENINVAIETT LWSIEWGIAELVNHPETQSKLRNELDVTVLGPGV	344
HtCYP76B1	-DNPNEFNHTNIKALFDLDFVAGTDTT SITIEWAMTELLRKP HIMS KAKEELEKVIKGS	331
AaCYP71AV1	-SAEFLP TSDNIKAIILDMFGAGTDTSSSTIEWAISELIKCPKAMEKVQAE L R KALNGKE	338
	: : : . : : : * .	
RrXp1A	-----AAIINEMVRMDPPQ-LSFLRFPTEDVEIGGVLIEAGSPIRFMIGAAN	470
PpP450cam	-----CEELLRRFSLVA--DGRILTSDYEFHGVQLKKGQDQILLPQMLSG	327
HsCYP19A1	DIKIDDIQKLVKMFENFIYESMRYQPVDLVMRK-ALEDDVIDGYPVKKGNTIILNIGRMH	402
BmP450-BM3	VPSYKQVQKLYVGMVLEALRLWPTAPAFSLYAKEDTVLGGEYPLEKGDELMLVLPQLH	362
HsCYP3A4	PPTYDVTLMQMEYLDVMVNETLRLFP-IAMRLERVCKKDVEINGMFIKPGVVMIPSYALH	402
HsCYP2D6	RPEMGDQAHMPYTTAVIHEVQRFGDIVPLGVTHMTSRDIEVQGFRIKPGTTLITNLSSVL	403
AtCYP73A5	QVTEPDLHKLPLYLQAVVKETLRLMAIPLLVPHMNLHDAKLAGYDIPAESKILVNAWWLA	404
HtCYP76B1	IVKEDDVLRLPYLSCIVKEVLRLLHPPSPLLLPRKVVTVQVELSGYTIAGTLVFNVAWAIG	391
AaCYP71AV1	KIHEEDIQELSYLNMVIKETLRLHPLPLVLPRECQPVNLAGYNIPNKTKLIVNVFAIN	398
	* * : :	
RrXp1A	RDPEVFDD-PDVFDHTRPP-----AASRNLSFGLGPHSCAGQIISRAEATTVFAV	519
PpP450cam	LDERE-----NACPM-----HVDFSRQKVSHTTFGHGSHLCLGQHLARREIIVTLKE	374
HsCYP19A1	RLE-FFPK-PNEFTLENFAKN-----VPYRYFQPFGGFPRCCAGKYIAMMMKAILVT	453
BmP450-BM3	RDKTIWGDVVEEFRPERFENPSA----IPQHAFKPFNGGQRACIGQQFALHEATLVLGM	417
HsCYP3A4	RDPKYWTE-PEKFLPERFSKKNKD---NIDPYIYTFPGSGPRNCIGMRFALMNMKLALIR	458
HsCYP2D6	KDEAVWEK-PFRFHPEHFLDAQGHFV--KPEAFLPFSAGRRACLGEPLARMEFLFFTS	459
AtCYP73A5	NNPNSWKK-PEEFRPERFEEESHVEANGNDFRYVPPFGVGRRCPGIILALPILGITIGR	463
HtCYP76B1	RDPTVWDD-SLEFKPQRFLES--RLDVRGHDFDLIFFGAGRRICPGIPLATRMVPIMLGS	448
AaCYP71AV1	RDPEYWKD-AEAFIPERFENS--SATVMGAEYELPFGAGRRMCPGAALGLANVQLPLAN	455
	* . * : * * : . :	
RrXp1A	LAERYERIELAEPTVAHND-----FA----RRYRKLPIVLS-----	552
PpP450cam	WLTRIPDFSIAPGAQIQHKSIGI-----VS-----GVQALPLVNDPATTKAV-----	415
HsCYP19A1	LLRRF-HVKTLQGGQCVESIQKI----HDLSLHPDETKNMLEMIFTPRNSDRCLH----	503
BmP450-BM3	MLKHF-DFEDHTNYELDIKETLTLKPEGFVVKAKSKKIPLGGIPSPSTEQSAKKVRRKAE	476
HsCYP3A4	VLQNF-SFKPKETQIPLK-----LSLGGLL---QPE-KPVVLKVESRDGT-----	499
HsCYP2D6	LLQHF-SFSVPTGQPRPSH-----HGVFALVSPSPYELCAVPR-----	497
AtCYP73A5	MVQNF-ELLPPPGQS--KVDTS-EKGGQS---LHLLNHSIIVMKPRNC-----	505
HtCYP76B1	LLNNF-DWKIDTKVPYDVLDMT-EKNGTTI---SKA-KPLCVVPIPLN-----	490
AaCYP71AV1	ILYHF-NWKLPNGVSYDQIDMT-ESSGATM---QRK-TELLLVPSF-----	495

**Figure 1.3:** Alignment of nine different cytochrome P450 enzymes: human CYP19A1 (aromatase), CYP3A4 and CYP2D6; bacterial P450-BM3, Xp1A, P450cam; plant CYP71AV1, CYP76B1, CYP73A5. Highlighted in red is the conserved cysteine ligand, in blue are the conserved sequences mentioned in the text. Underlined in green are the conserved eukaryotic N-terminal residues. For P450 BM3, only the heme domain has been included.

Like the variation seen at the primary sequence level, the substrate-recognition sites and the signal peptides for the targeting/incorporation of these proteins into the membrane are also very variable regions in P450 structures. Despite this, an overall

tertiary structure is preserved among all P450s, with alpha and beta domains, the former containing the catalytic centre, the latter forming the substrate access channel, with the substrate recognition site, assembled to form a conical prism (figure 1.4).



**Figure 1.4:** Top-Schematic representation of the general secondary and tertiary structure observed for cytochrome P450 enzymes (scheme from Werck-Reichhart and Feyersen (15)). Bottom-Overlapped ribbon structures of four CYPs: CYP3A5, CYP2B4, CYP2C5, CYP2C9 (figure from Bak *et al.* (17)), demonstrating how the overall tertiary conical structure is conserved across different CYPs. Depicted in pink is the heme in the catalytic site, in orange is the heme-binding loop.

In 1987, Poulos *et al.* solved the first P450 structure, that of a camphor hydroxylase, P450cam, from *P. putida* (PDB ID: 2CPP) (18). According to the CYPED Cytochrome P450 Engineering Database (<https://biocatnet.de/>, (19)), a platform created by Prof.

Pleiss' group at the University of Stuttgart, 595 pdb entries, corresponding to 595 different CYP protein structures, have been deposited so far. In some cases, such as for CYP2B4, CYP3A4, CYP55A1, CYP101A1, multiple isoforms are present for each CYP, due to mutated residues or for the presence of a substrate or an inhibitor in the catalytic site.

### **Nomenclature**

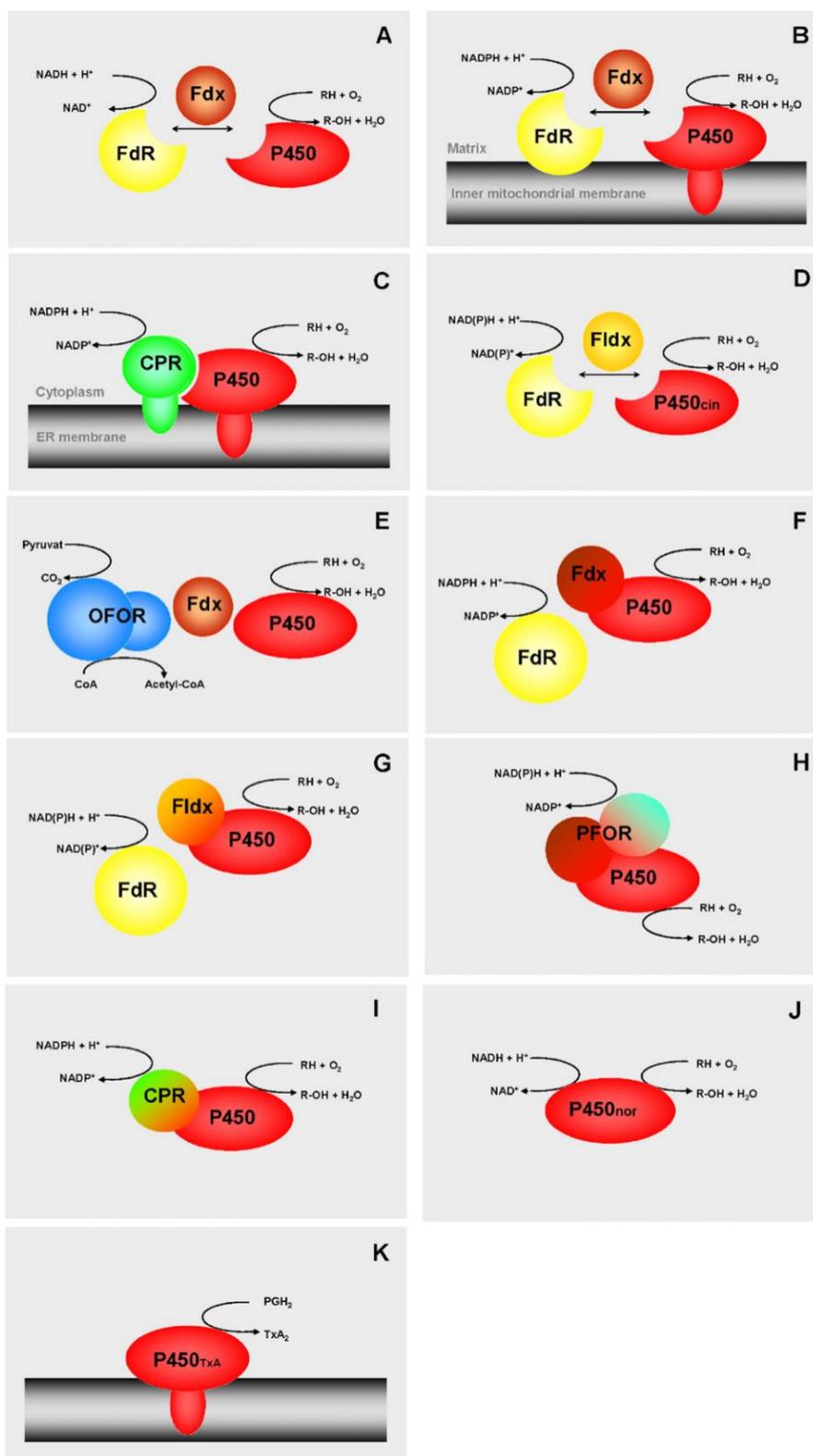
The nomenclature of P450 enzymes is based on a triple alphanumeric code, where the abbreviation 'CYP' is followed by the family number, which groups all the P450s with > 40 % sequence identity; a letter, which identifies the subfamily of CYPs with > 55 % sequence identity, and the gene number, e.g. CYP73A5. There are a few exceptions to the rule, such as for example *Bacillus megaterium* CYP102A1 and *P. putida* CYP101, more commonly mentioned in studies and reports as P450 BM-3 and P450cam, respectively. The family and subfamily numbering system depends on the chronological order of gene discovery. Animals, lower eukaryotes and plants were allocated the first 100 families, with CYP1-49 corresponding to animals, CYP51-69 to fungi and CYP71-99 to plants. Bacterial CYPs span from CYP101 onward. Soon, in many cases, the family range was exceeded, therefore additional 200 'blank spaces' were added for each group of organisms, for the classification of newly-found families: in this way bacterial sequences spanned from CYP101 to CYP299, animal CYPs from CYP301 to CYP499, lower eukaryotes from CYP501 to CYP699 and plants from CYP701 to CYP999. In the case of lower eukaryotes, all the digits were again filled, and for this reason new allocations for each group were defined, similarly as before, with numbers up to CYP9999 (16). A more in-depth phylogenetic analysis of the CYP sequences, encouraged by widening access to genomic data (for example, the complete *Arabidopsis thaliana* genome sequence, published by the Arabidopsis Genome Initiative in 2000 (20)), led to the addition of a further classification layer, where CYP families are subdivided into clans. By comparing the genomic data available for rice (*Oryza sativa*) and Arabidopsis, with a total of 727 CYP gene sequences, Nelson *et al.*, defined ten different plant clans, some consisting of

numerous families, such as CYP71, CYP72, CYP85 and CYP86, named after the lower family number present in the grouping (21).

## Organisation

While all known prokaryotic P450 enzymes are cytosolic, eukaryotic CYPs are mainly bound to membranes, either on the outer side of the endoplasmic reticulum (ER), as in the case of the majority of plant P450s, or associated with the inner mitochondrial membranes.

The activity of CYP enzymes requires the presence of redox partners, which allow the creation of a redox potential gradient for the efficient shuttling of electrons from the donor, in most cases NADH or NADPH, to the heme. For this reason, a cytochrome P450 reductase, with bound flavin adenine di/mononucleotides or FeS clusters, is usually co-localised and co-expressed with the CYP (a more detailed overview of the redox partners can be found in the next section). The redox chain may be laid out as two/three independent soluble modules (class I and III bacterial systems) or associated via N-terminal anchors to the membranes (microsomal systems) or completely fused as a single protein (see P450-reductase arrangements in figure 1.5). Examples of the latter are the *BmP450* BM-3 and P450-PFOR systems from *Rhodococcus* and *Burkholderia* (22). Another example is XplA (CYP177A1) from *Rhodococcus rhodochrous*, where the heme domain is fused to a flavodoxin domain. XplA is encoded in the same gene cluster as the FAD-bound reductase XplB (23). Exceptions are given by some independent P450s which do not require a reductase for the catalytic activity. Examples of the latter case are CYPs that exploit the reversed-peroxide shunt (see section 1.3), using H<sub>2</sub>O<sub>2</sub> as a 'trigger' for the direct formation of one of the final reactive intermediates of the catalytic cycle, skipping the general sequence of reductions and oxidations (24). Other examples of autonomous P450 enzymes are the fungal P450<sub>nor</sub> (J in figure 1.5) and human P450<sub>TXAS</sub> (K in figure 1.5) which perform the denitrification of nitric oxide and the conversion of prostaglandins to thromboxanes, respectively.



**Figure 1.5:** Representation of the different P450-redox partner configurations. (A) bacterial system (class I), (B) mitochondrial system (class I), (C) microsomal system (class II), (D) bacterial system (class III), (E) bacterial thermophilic system (class IV), (F) bacterial fusion system (class V), (G) class VI, bacterial fusion system (class VI), (H) bacterial fusion system (class VII), (I) bacterial fusion system (class VIII), (J) eukaryotic P450<sub>nor</sub> (class IX), (K) P450<sub>TxA</sub>, an independent eukaryotic system (figure from Hannemann *et al.* (22)).

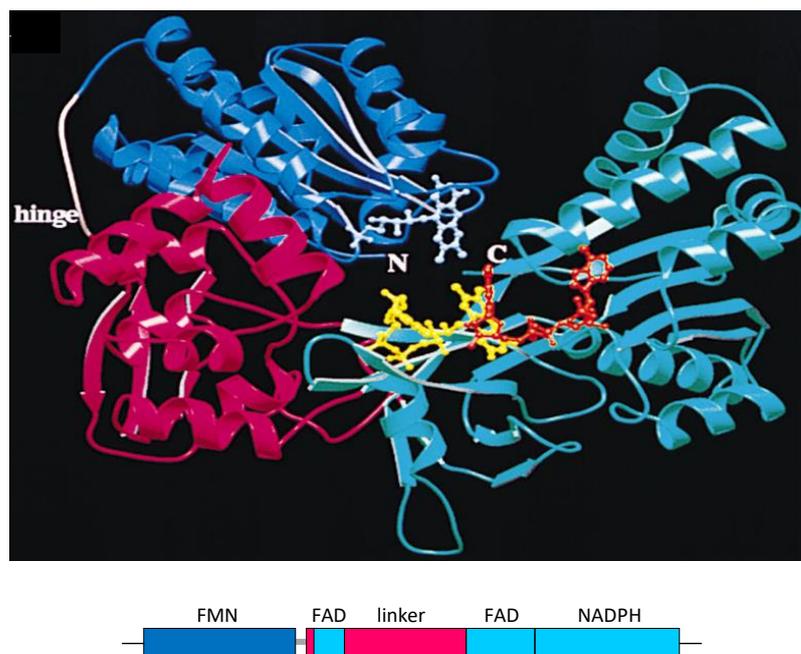
The activity of P450 enzymes can also be affected, in a positive or negative way (no general rule is applicable) by the presence of microsomal  $\text{cyt}_{b5}$ , which can act as a mediator of the electron transport between the P450 and the reductase domain. Zhang and colleagues investigated the effect of  $\text{cyt}_{b5}$  on CYP2B4 catalysis, in the presence or absence of the natural reductase partner. This study revealed an increased product formation when  $\text{cyt}_{b5}$  was present in a 1:1 ratio with the CYP, when compared to the CYP-reductase conventional pairing. In addition, the same research group discovered a competition between  $\text{cyt}_{b5}$  and the reductase for the binding to a specific site on the CYP when both the electron donors are present (25).

### **Cytochrome P450 redox partners**

The electron-shuttling in cytochrome P450 catalysis relies on the presence of a redox chain, constituted, in the case of microsomal CYPs, by a cytochrome P450 reductase (abbreviated to CPR or POR or CYPOR). This reductase partner, co-localised in the ER-membrane, is a diflavin enzyme, presenting three binding sites: two for the flavin cofactors (flavin mononucleotide, FMN, and flavin adenine dinucleotide, FAD) and one for the reduced nicotinamide adenine dinucleotide phosphate (NADPH). Vermilion and Coon characterised the purified CPR through photochemical reduction. Comparisons between the native reductase and an FMN-depleted reductase were carried out in order to evaluate the respective contributions of the two flavin cofactors in the redox chain. It was concluded that FAD is the initial acceptor of reducing equivalents from NADPH, which are then passed on to the FMN. Subsequently, FMN fulfills the stepwise one-electron transfer to the final acceptor, the P450 heme (26).

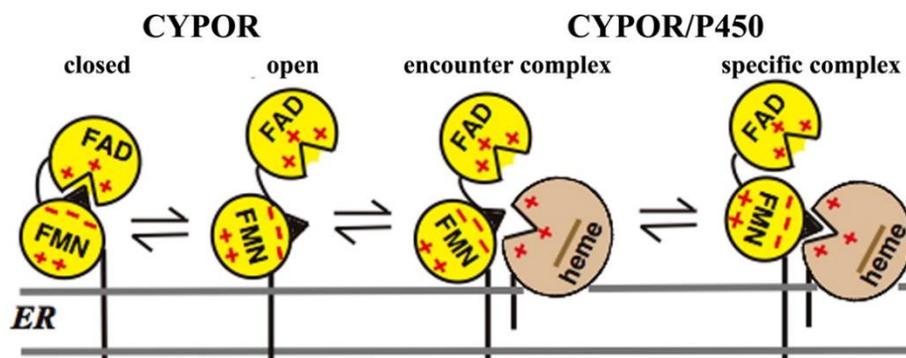
The elucidation of the 3D structure of a rat microsomal CPR by Wang *et al.*, allowed further understanding of the spatial arrangement of the domains constituting the redox chain. The authors identified an N-terminal hydrophobic domain, constituting the anchor to the ER membrane, followed by the FMN-binding domain and a linker region of approximately 150 amino acids. This region, consisting mainly of  $\alpha$ -helices was assigned the role of ensuring the optimal orientation and distance between the flavin domains, which, upon electron acceptance and transfer, change spatial

configuration. The FAD- and NADPH-binding domains were localised at the C-terminus of the peptide chain (27) (figure 1.6).



**Figure 1.6:** Ribbon representation of the crystal structure of rat cytochrome P450 reductase (PDB ID: 1AMO). From N-terminus to C-terminus: the FMN-binding domain is coloured in dark blue (bound FMN in light blue), linker region in fuchsia, FAD and NADPH-binding domains in light blue (bound FAD and NADPH in yellow and red, respectively) (figure from Wang et al. (27)). At the bottom, a linear representation shows the arrangement of the FMN/FAD/NADPH-binding domains, with the same colour code as for the above ribbon representation.

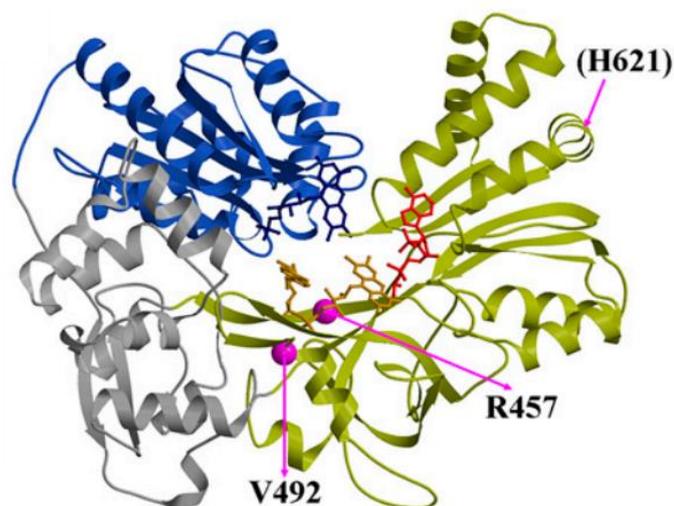
Figure 1.7 represents the proposed arrangement and interaction among subunits: the N-terminal domain of the CPR, with the bound FMN, is inserted in ER membrane. In the final stage, when the FMN domain has to transfer the reducing equivalents to the P450 heme, the protein configuration changes, thanks to the flexible central linker, from a 'close' to an 'open' setting (28).



**Figure 1.7:** Interaction between the domains (N-terminal FMN and C-terminal FAD, then FMN with P450 heme) involved in the electron flux, from the donor (NADPH) to the final acceptor (CYP). CYPOR = CPR (figure from Iyanagi *et al.* (29)).

Sequence alignments between the characterised rat CPR and other flavoenzymes, such as rat neuronal nitric-oxide synthase, bacterial *B. megaterium* P450 BM-3, *Escherichia coli* sulfite reductase, *Desulfovibrio vulgaris* flavodoxin and the plant *Spinacia oleracea* ferredoxin-NADP<sup>+</sup> reductase, demonstrated a high level of sequence conservation (up to 58 %), especially at the cofactor-binding regions and for most of the secondary structure elements (27). This reflects the promiscuity of the CPR, which, in mammals as well as in other systems, interacts with a multitude of highly diversified CYPs. Radioimmunoassays performed on purified CYPs and CPRs from rats, collected from samples treated with phenobarbital and  $\beta$ -naphthoflavone (NF), revealed a ratio of CYP to CPR of 15:1 for the untreated controls and of up to 21:1 for the NF-treated samples (30).

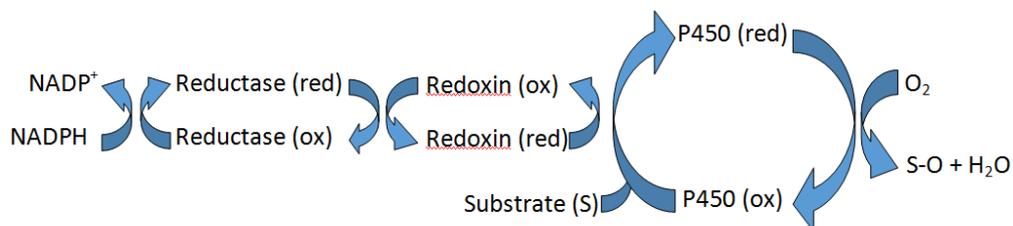
Humans possess a single CYP reductase gene, which displays a > 90 % sequence identity with the rat CPR. Therefore, the information gathered by Wang *et al.* was an excellent basis to build on for the structural characterisation of the human CPR. As expected, Xia and colleagues discovered a high degree of similarity between the conformation of the two proteins, in terms of secondary structures and cofactor binding domains. The resulting human CPR structure, solved at a higher resolution than for the rat CPR (at 1.75 Å versus 2.6 Å, respectively), is reported in figure 1.8.



**Figure 1.8:** Ribbon representation of human cytochrome P450 reductase. From N- to C- terminus, the FMN-binding domain is represented in blue (same colour for the bound FMN), the linker region is in grey, whereas the FAD- and NADPH-binding domains are in yellow (bound FAD in orange and bound NADPH in red). The highlighted amino acids are mutation sites, which affect the stability of the FAD-binding domain. H621, in brackets, is a residue present in the rat CPR but not in the human CPR (figure from Xia *et al.* (31)).

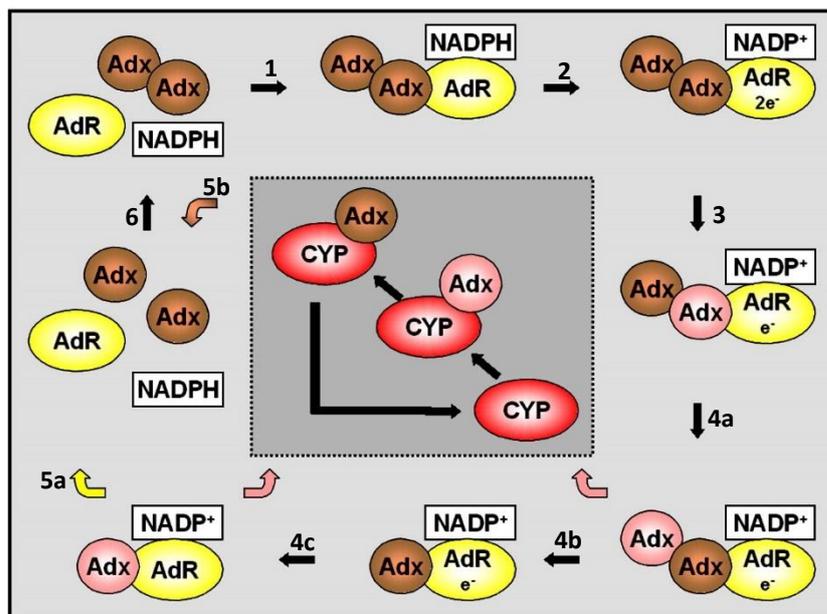
Multiple studies reported that cytochrome P450 enzymes are not the only partners for CPR. In fact, since the first reports published in the late '60s about the CYP/CPR complex (32), it has been discovered that CPR can also support the activities of heme oxygenase and squalene epoxidase (33).

In the case of bacterial and mitochondrial eukaryotic CYPs (class I P450s), the redox partner system may consist of an FAD-containing reductase domain, which receives reducing equivalents from either NADH or NADPH, shuttles them to a partner ferredoxin (which contains FeS clusters) or to an FMN-binding flavodoxin and, finally, to the heme partner. In bacteria all the modules of the electron chain are soluble, whereas in mitochondrial systems the first electron recipient is membrane-bound, as is the CYP partner, and the ferredoxin intermediate partner is soluble. The most well-described case of a class I cytochrome P450 is P450cam (CYP101) from *P. putida*, a camphor hydroxylase discovered in 1968 (12). In P450cam, the electrons from NADH are shuttled through an FAD-binding putidaredoxin reductase, then to an FMN-containing putidaredoxin and finally to the P450 heme (34) (figure 1.9).



**Figure 1.9:** Representation of the electron transport chain in the P450cam from *P. putida*. Electrons from NADPH are shuttled through the redox partner (first the reductase then the redoxin domain) and, ultimately, to the P450 heme.

In mammals, adrenodoxin (AdX) is the main redoxin partner. The structure of AdX was firstly solved by Pikuleva *et al.* (35), and several studies have since offered suggestions for the exact mechanism of electron transport from NADPH to the heme. As proposed by Beilke *et al.* (36), the reduction of the P450 substrate depends on a four-step process, which is initiated by the binding of NADPH to adrenodoxin reductase (AdR), followed by the coupling with a dimer formed by two molecules of AdX. Once reduced, the Adx monomers detach from the tetramer one by one and become re-oxidised after the interaction with the CYP. Finally, the two oxidised AdX dimerise again for a new cycle of P450 reaction (figure 1.10).



**Figure 1.10:** Electron transport chain in the mammalian AdR/AdX-CYP system (figure modified from Ewen *et al.* (37)). The multistep process starts with the binding of NADPH to Adrenodoxin Reductase (AdR) (step 1); two electrons are donated from NADPH first to AdR (step 2) and then, in a stepwise manner, to each of the Adrenodoxin (Adx) molecules present in the system (steps 3-4c). Each reduced Adx donates an electron to the CYP (inner square). The free, re-oxidised AdR and Adx molecules (5-6) are now ready for a new cycle.

Schiffler and Bernhardt carried out a comparative study between *P. putida* P450cam and human CYP11A1, CYP11B1, CYP11B2 redox chains. Through amino acid sequence alignments, EPR studies and topology examination, the authors demonstrated a high degree of similarity between the redoxin partners in the two systems (putidaredoxin for P450cam and adrenodoxin for the human CYP11s) (38).

With the aim of enhancing CYP activities or assigning new substrate/product specificities many researchers have manipulated the CYP/CPR system in different ways. As an example, Neunzig and colleagues heterologously expressed a set of human P450 enzymes in yeast and co-expressed yeast or plant redox partners. A comparison of these mixed systems with the human CYP/CPR system showed that CYP2D6 activity was best supported by non-human partners (39). Also mutagenesis of the redox partner can help increase the efficiency of the electron transport to the CYP, as demonstrated by Schiffler *et al.* In this study, the group succeeded in the significant improvement of the interaction between a bovine adrenodoxin and different CYP11 enzymes involved in the synthesis of hormones (40).

### 1.3 Catalytic cycle of cytochromes P450

Cytochrome P450 enzymes are generally described as monooxygenases, as they cleave molecular oxygen while reducing equivalents, in a stepwise manner, are shuttled to the active site through the associated redox chain (as described in the previous section), leading to the final, concomitant release of the oxygenated product and water.



The P450 catalytic cycle has been described extensively over the years. Generally, it starts with the access of the substrate in the active site, causing the displacement of the water ligand and a structural rearrangement (step 1 in figure 1.11). A first electron is donated by NAD(P)H through the reductase partner and the heme shifts from the low to the high spin state (step 2). An oxidative step with dioxygen then occurs (step 3), followed by the second electron donation (step 4). Two consecutive protonation steps yield first a ferric hydroperoxo-complex (step 5) then the highly reactive 'Compound I' (step 6), which induces the rearrangement of the substrate, leading to the release of the final product (steps 7 and 8). The dissociation of the oxygenated compound allows the system to return to the initial oxidated stage, with bound water, ready to catalyse a new reaction (41-44).



they observed that more than 90 % of the oxygen consumption of the system, when CYP3A4 was incubated with testosterone, was associated with uncoupling reactions (mainly to the oxidase shunt, with release of water from Compound I) (43).

Cytochromes P450 are able to catalyse a vast array of reactions, from monooxygenations to hydroxylations, epoxidations, N-/S-/O- dealkylations to N-demethylations, isomerisations, dimerisations, reductions, dehalogenations and deaminations (41, 42). For this reason, CYP enzymes are very interesting targets for the production of high-value chemicals for the food, pharma and cosmetic sectors. Additionally, CYPs can be exploited for pharmacological ADME/Tox studies due to their involvement in drug metabolism (mainly CYP3A4 and CYP2D6 in human), as also for agronomical and bioremediation purposes.

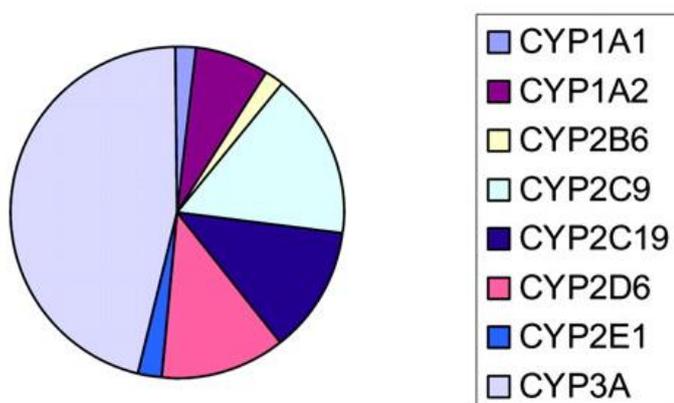
## 1.4 Human Cytochromes P450

As confirmed by the outcomes from the Human Genome Project almost 15 years ago (45), the human genome contains 57 genes encoding cytochromes P450, subdivided into 18 families (46): many are expressed in the liver, where they are involved in the metabolism of xenobiotics, leading to either the detoxification or the bio-activation of target compounds. P450 enzymes are also located in the intestines (mainly CYP3A, CYP2C (47)), in the lungs (CYP1A1, CYP1B1, CYP2F1, CYP2S2, CYP4B1), kidneys and in the brain (CYP2E1, CYP2D6, CYP2B6 (48)). In most cases, human CYPs are microsomal, with the exception of seven CYPs, located in the mitochondria. Moreover, human CYPs display broad substrate specificity and can catalyse multiple reactions, such as the catabolism/activation of xenobiotics and the synthesis of endobiotics (steroid hormones, fatty acids, eicosanoids and vitamins (table 1.2)). An important example is CYP19A1, also named aromatase, responsible for the synthesis of estrogens (49). Mutations, deficiency or excess of CYP19A1 can lead to hormonal imbalance, ultimately causing endometrial cancer, metastatic breast cancer, virilisation (50-52).

**Table 1.2:** Subdivision of the human CYPs based on major substrate specificity. N.A.= not available, referred to orphan CYPs (classification from Guengerich (53)).

Sterols	Xenobiotics	Fatty Acids	Eicosanoids	Vitamins	N.A.
CYP1B1	CYP1A1	CYP2J2	CYP4F2	CYP2R1	CYP2A7
CYP7A1	CYP1A2	CYP2U1	CYP4F3	CYP24A1	CYP2S1
CYP7B1	CYP2A6	CYP4A11	CYP4F8	CYP26A1	CYP2W1
CYP8B1	CYP2A13	CYP4B1	CYP5A1	CYP26B1	CYP4A22
CYP11A1	CYP2B6	CYP4F11	CYP8A1	CYP26C1	CYP4X1
CYP11B1	CYP2C8	CYP4F12		CYP27B1	CYP4Z1
CYP11B2	CYP2C9	CYP4F22		CYP27C1	CYP20A1
CYP17A1	CYP2C18	CYP4V2			
CYP19A1	CYP2C19				
CYP21A2	CYP2D6				
CYP27A1	CYP2E1				
CYP39A1	CYP2F1				
CYP46A1	CYP3A4				
CYP51A1	CYP3A5				
	CYP3A7				
	CYP3A43				

As reported by Williams *et al.*, more than 75 % of the administered drugs are metabolised by P450 enzymes, followed by UDP-glucuronosyltransferases and esterases (54). Among the CYPs highlighted in this study, CYP3A4 was the biggest contributor to the bioconversion of drugs (figure 1.12).

**Figure 1.12:** Group of P450 enzymes with relative contribution to drug metabolism. CYP3A4 was identified as the biggest contributor, followed by CYP2C9, CYP2D6 and CYP2C19 (figure from Williams *et al.* (54)).

In addition to its activity towards pharmaceuticals, such as the antineoplastic ellipticine (55) and the anticancer prodrugs cyclophosphamide and ifosfamide (56), CYP3A4 has been reported as active towards natural toxins, such as the aflatoxin B1 from *Aspergillus* (57); and carcinogens, such as the polycyclic hydrocarbon benzo- $\alpha$ -pyrene (58).

Human CYPs have also been tested towards environmental pollutants, such as insecticides, herbicides and volatile organics. Afterwards, plant systems such as rice, tobacco and *Arabidopsis* have been engineered to express these reactive human CYPs (mainly CYP2E1 and CYP1A1, CYP2B6, and CYP2C19) to introduce resistance to the toxic compounds (59-62).

Human cytochrome P450 enzymes can also be active towards endogenous substrates. The already-mentioned hepatic and intestinal CYP3A4 displays a very broad activity (63), as it is also able to hydroxylate cholesterol (64) and a range of endogenous steroid hormones, such as progesterone and androstenedione (65-67). It has been observed that the expression of human CYP4A2 in rat led to an increased amount of renal 20-hydroxyeicosatetraenoic acid (20-HETE). This arachidonic acid regulates the contractions of smooth muscles, affecting the vascular flow, and resulting ultimately in higher blood pressure and hypertension. When *N*-hydroxy-*N'*-(4-butyl-2-methylphenyl)-formamidine (HET0016, a known inhibitor of 20-HETE synthesis), was supplied, the effect was reversed (68). Similar results were obtained also with CYP4A11 and CYP4A12, the expression of which were observed as strictly correlated to the overproduction of 20-HETE (69, 70).

Genetic variations of the human P450 enzymes can result in deleterious phenotypes such as abnormalities in the rate of absorption, delivery or excretion of drugs and/or diseases. As highlighted on the CypAllele website (curated by Sim and Ingelman-Sundberg from the Karolinska Institutet (71)), a particular case is that of CYP2D6. This P450 enzyme has been found in numerous (more than 90) variants, with varying frequencies across the different ethnicities (72). As an example, *CYP2D6*\*17, a common variant found in Black African individuals, leads to a decreased CYP2D6 activity. CYP2D6 expressed in the brain has a documented neuroprotective activity

against toxins (73). Individuals displaying decreased or deleted CYP2D6 activity, lack this protection and therefore are more prone to neurodegenerative diseases. This was demonstrated in a genetic study where the levels of CYP2D6 were compared between Parkinson's patients and healthy controls (74, 75). On the opposite side, individuals expressing multiple copies of CYP2D6, exhibit an excessively fast drug metabolism, with resulting toxicity, as demonstrated in the case of codeine. Two specific variants of the *CYP2C19* gene, *CYP2C19\*2* and *CYP2C19\*3*, cause the complete inactivation of the catalyst. For this reason, the subjects, presenting a 'poor' metabolism towards the substrates of CYP2C19, such as drugs, have to undertake alternative therapeutic treatments (either in terms of active compound or in dosage) as compared to 'normal' metabolisers (76). Another important aspect to take into account is the effect of the inhibition of drug-metabolising P450s: by blocking these detoxifying activities, the drug compound, which remains unaltered, accumulates, with resulting toxicity. This can occur especially when multiple therapies are administered at once. As an example, numerous reports showed that selective serotonin reuptake inhibitors (SSRI) antidepressants compete with a range of therapeutic substrates, such as the antidepressant imipramine and the anxiolytic alprazolam, for the binding to the CYP2D6 active site, therefore blocking the related activity and decreasing the final clearance of the compounds (77-79). SSRIs can also inhibit CYP3A4 catalysis. A well-documented CYP3A4 inhibitor is the antifungal ketoconazole, which, leads, as reported by Gomez *et al.*, to the overaccumulation of the immunosuppressant cyclosporine (80). As reported by Koudriakova and co-workers, also the HIV protease inhibitor ritonavir competes with the CYP3A4 substrate. The co-administration of ritonavir, and resulting inhibition of CYP3A4-mediated activity, was seen by Kempf *et al.* as an option to induce the bioaccumulation of concomitantly used anti-HIV medications and potentiate their pharmacological effect (81). From the pharmaceutical point of view, considering the involvement in the activation or detoxification of drugs, human cytochrome P450 enzymes constitute a precious source of information for pharmaceutical companies, in toxicological studies as well as in the design of novel therapeutics with diversified structures or modified mode of action. Furthermore, screening assays for the determination of allelic variants of P450 drug metabolising enzymes (CYP3A4, CYP2D6, CYP2C9, etc.) should

be of particular help in the design of personalised therapies, whereby dysfunctional catalysts are identified.

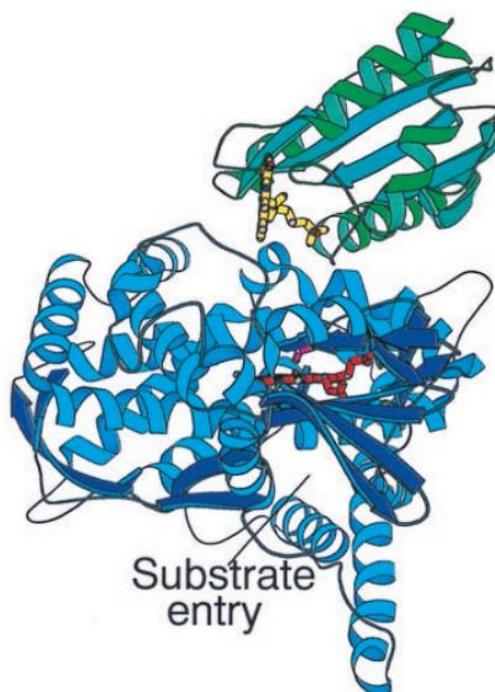
## 1.5 Bacterial Cytochromes P450

Bacterial CYPs are soluble enzymes and for this reason the characterisation of these systems has been easier to complete, when compared to the membrane-bound eukaryotic enzymes. Bacterial cytochrome P450s catalyse a wide array of reactions, such as the synthesis of pharmacologically active compounds (as secondary metabolites), the breakdown of xenobiotics and the metabolism of fatty acids. The two best characterised bacterial P450 systems, which have been mentioned many times throughout the previous sections, are *B. megaterium* P450 BM-3 (CYP102A1) and *P. putida* P450cam (CYP101). Characterised bacterial CYP activities are listed in table 1.3.

**Table 1.3:** Examples of known bacterial P450-mediated activities (information from Kelly and Kelly (82)).

Organism	CYP	Function
<i>Pseudomas putida</i>	CYP101A1 (P450cam)	catabolism of camphor
<i>Bacillus megaterium</i>	CYP102A1 (P450-BM3)	fatty acid catabolism
<i>Streptomyces avermitilis</i>	CYP105P1, CYP105D6	filipin biosynthesis/antifungal
<i>Saccharopolyspora erythraea</i>	CYP107A1	erythromycin biosynthesis/antibacterial
<i>Streptomyces lavendulae</i>	CYP107N1, CYP160A1, CYP105F1	mitomycin c biosynthesis/antitumour
<i>Streptomyces hygroscopicus</i>	CYP122A2	rapamycin biosynthesis/immunosuppressant/anti-ageing?
<i>Mycobacterium tuberculosis</i>	CYP125A1	catabolism of cholesterol
<i>Jeotgalicoccus</i> sp. ATCC 8456	CYP152A3	decarboxylation of fatty acid
<i>Streptomyces fradiae</i>	CYP105L, CYP113B1, CYP154B1	tylosin biosynthesis/veterinary antibacterial
<i>Streptomyces noursei</i>	CYP105H1, CYP161A1	nystatin biosynthesis/antifungal
<i>Streptomyces spheroides</i>	CYP163A1	novobiocin biosynthesis/antibacterial
<i>Sorangium cellulosum</i>	CYP167A1	epothilone biosynthesis/antitumour
<i>Streptomyces avermitilis</i>	CYP171A1	ivermectin biosynthesis/anthelmintic and insecticide

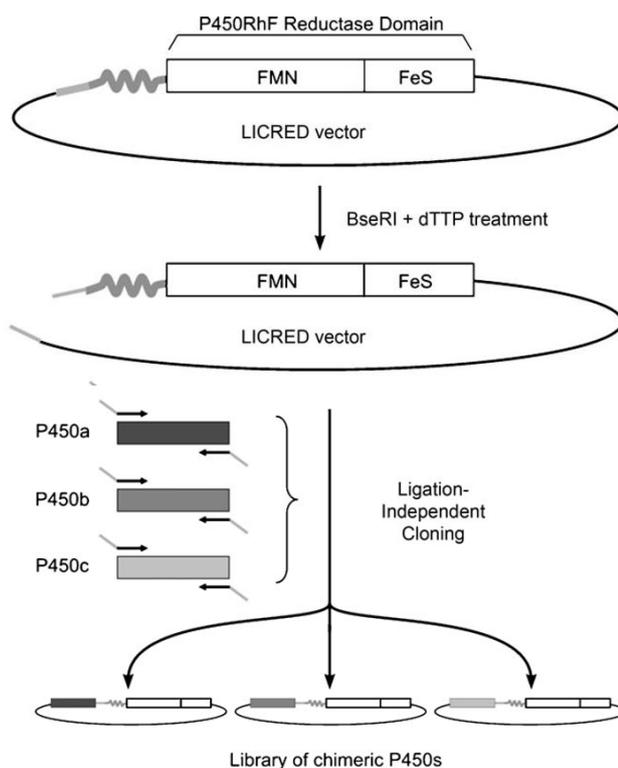
The P450 BM-3 fusion was first described in 1986 by Narhi and Fulco, who were studying the hydroxylation of fatty acids in *B. megaterium* ATCC 14581. On the basis of previous reports demonstrating that the hydroxylating activity and the expression of the responsible enzyme are highly induced by barbiturates, the two researchers used phenobarbital to induce the overexpression of the enzyme for purification and biochemical characterisation. Upon SDS-PAGE analysis, the isolated enzyme corresponded, on the gel, to a single 119,000 kDa band, which was double the size of previously-described CYPs. Trypsin digestion revealed an N-terminal 54 kDa heme domain and a 65 kDa C-terminal reductase domain. Differently from other bacterial system, the redox partner had the features of eukaryotic CPRs, as it presented bound FAD and FMN, as verified using spectrophotometrical observations (83, 84). This soluble fusion system was active towards long-chain fatty acids, using preferentially NADPH (rather than NADH) as an electron donor and, displaying the highest catalytic activity in the P450 world, with a turnover number of 280 s<sup>-1</sup> (85). The structure of the P450 BM-3 heme domain and FMN-binding domain was elucidated a few years later, by Sevrioukova *et al.* (figure 1.13).



**Figure 1.13:** Structure of *B. megaterium* P450 BM-3 (lacking the FAD-binding domain), as elucidated by Sevrioukova *et al.* (86). From N-terminus to C-terminus: the P450 heme domain is represented in blue, the heme in red, the FMN-binding domain in green (bound FMN in yellow).

Two other interesting bacterial P450 systems are XplA (CYP177A1) and P450RhF (CYP116B2), both from *Rhodococcus*. The XplA enzyme, which has only been found at military sites or manufacturing areas contaminated with explosives, is a successful example of the application of microbial CYPs for biotechnological purposes. This enzyme, with a fused flavodoxin domain, is key for the survival of the microbes in soils heavily contaminated with the toxic explosive hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX). In fact, XplA catalyses the denitration and cleavage of the aromatic ring of the RDX molecule, enabling the bacteria to use it as a source of nitrogen for growth (87). Transgenic plants expressing XplA were able to biotransform and retain the RDX taken up from the soil. This novel technology for bioremediation, currently under field trials in the United States, constitutes a promising sustainable tool to clean up explosive-contaminated lands (23, 88).

Roberts and colleagues isolated P450RhF from *Rhodococcus* sp. strain NCIMB 9784. This enzyme was another example of a fused P450 system, presenting an N-terminal heme domain connected through a 16 amino acid-linker to a C-terminal redox partner, composed of an FMN-binding domain, followed by an NADPH-binding domain and a 2Fe-2S cluster. Activity screenings demonstrated that this CYP is able to catalyse the formation of 7-hydroxycoumarin by dealkylating 7-ethoxycoumarin (89). Since its discovery, this system has been exploited widely for the creation of chimeras, whereby P450 heme domains are connected to the Rhf-reductase domain (RhfRed = C-terminal section of P450RhF). This is the case, for example, with the LICRED platform (figure 1.14, 90) developed by Dr. Sabbadin in our laboratory, which has been validated with P450cam and XplA and subsequently tested also with *Nocardia farcinica* and human CYPs (91).



**Figure 1.14:** Illustration of the LICRED drop-in platform developed by Dr. Sabbadin. The construct consists of a C-terminal RhfRed domain (from P450RhF). The P450 heme targets can be cloned at the N-terminus by means of LIC-cloning (figure from Sabbadin *et al.* (90)).

Zhang and co-workers created three fusions of *Micromonospora griseorubida* P450MycG with RhfRed, with the spinach Fdx/FdxR redox system as well as with the whole P450RhF. Interestingly, the three fusions yielded final products that were different from the expected native products, mycinamicin antibiotics (92). Another recent example is that of Makino and co-workers, who engineered CYP-RhFRed fusion constructs in order to characterise the function of different CYP110 enzymes from cyanobacteria, identifying a particularly broad substrate specificity for *Nostoc* CYP110E1 (93).

The model bacterial system *Escherichia coli* does not have any cytochromes P450 and it has been routinely employed over the years for the heterologous expression of cytochrome P450 enzymes, in alternative to other bacteria (*B. megaterium*, *P. putida*), yeasts (*Pichia pastoris*, *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae*) and tobacco plants.

Numerous studies carried out by members of the Bernhardt group exploited the great potential of microbes as factories for the production of high-value chemicals. Among the reported successful processes are the introduction into *E. coli* of CYP106A2 from *B. megaterium* ATCC 13368, coupled with a bovine redox partner and an electron recycling system, for the regioselective hydroxylation of 3-oxysteroids (94). Further substrate screens with the same CYP complex, revealed novel hydroxylating activities towards diterpenes and triterpenes and 3-hydroxysteroids, such as pregnenolone and dehydroepiandrosterone (DHEA), in *B. megaterium* whole cell incubations. These results were particularly important from the pharmaceutical point of view, as hydroxylated steroids in the human organism display multiple effects, such as antiinflammatory, neuroprotective and immunoregulatory effects, as in the case of the obtained 7 $\beta$ -hydroxy-DHEA (95). It has been demonstrated that CYP106A2, as well as CYP106A1, in *B. megaterium* (strains MS941 and DSM319) whole cell assays, are also able to hydroxylate diterpenes and triterpenes (96, 97). More recently, the plethora of substrates for these two CYP106A has been further expanded, with the discovery, beyond the already characterised hydroxylations at the 6 $\beta$ , 7 $\beta$ , 9 $\alpha$ , and 15 $\beta$ -positions, of a novel 11-oxidase activity (98). On the same line, Bracco and colleagues generated mixed systems consisting of a *N. farcinica* CYP154C5, *P. putida* redox partners and *E. coli* for expression and activity assays towards a series of steroids. It has been discovered that this CYP was able to catalyse, differently from CYP106A1/A2, the 16 $\alpha$  hydroxylation of testosterone and all the other screened substrates, with higher efficiency than previous reports with P450 enzymes from *Rhodococcus* or from *Streptomyces* (99). The highly selective hydroxylations, of which some examples have just been described here, are particularly attractive for pharmaceutical companies, as this type of molecular functionalisation can directly affect the metabolism and bioavailability of drugs.

Researchers have created several mutants of known systems in order to direct catalysis towards desired products. As an example (among the multitude of similar protein engineering studies), Peters *et al.* successfully created, by means of directed evolution and site-directed mutagenesis, a mutant of P450 BM-3, enabling a novel

functionality: the stereo- and regio-specific hydroxylation of short and medium length alkanes, such as propane, heptane, octane (100).

## 1.6 Plant Cytochromes P450

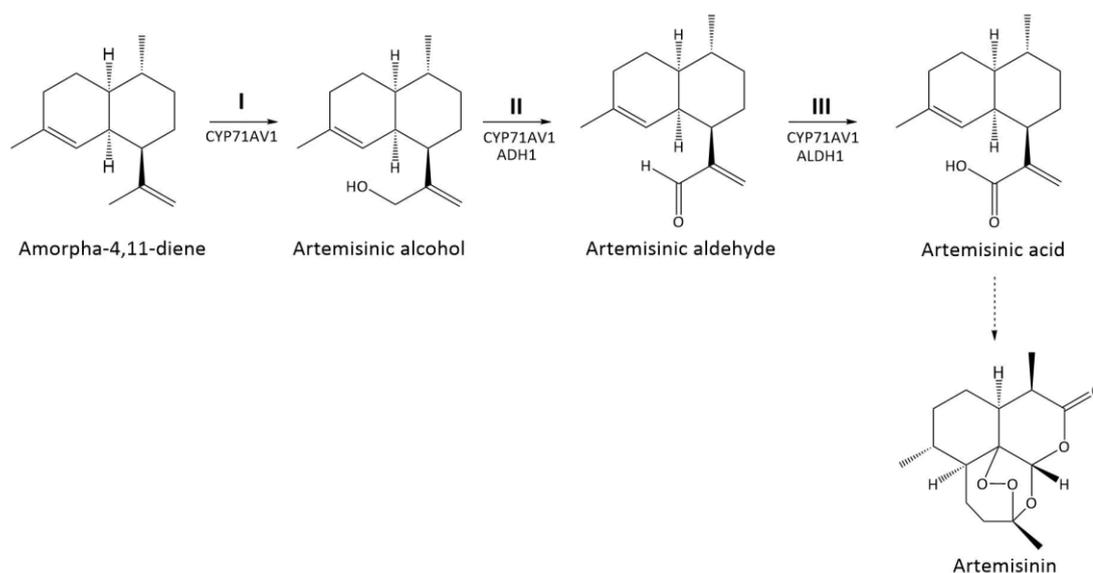
The most numerous group of P450 genes can be found in plants, with, as reported in section 1.1, more than 13,000 sequences annotated so far. Not counting the pseudogenes, Arabidopsis alone contains 244 cytochrome P450s, whereas rice (*O. sativa*) presents 350 CYP genes. These very high numbers, when compared to other living systems, reflect the importance of this group of enzymes, which, according to evolution and adaptation, have duplicated and mutated, generating different isoforms among the same gene family with amplified or diversified activities (101). Examples of this case are the numerous genes, 54, grouped in the CYP71 family, the largest among all plant P450 families, of which only a few have been identified as participants at different steps in the camalexin-based plant defence mechanism (17).

The activity of CYPs in plants spans numerous biosynthetic processes and metabolic pathways, leading to fundamental compounds for plant life. Among these are signalling and defence molecules (e.g. oxylipins, phytoalexins, glucosinolates, alkaloids, cyanogenic glucosides), developmental hormones (gibberellins, brassinosteroids, abscisic acid, cytokinins) or structural components (lignins, cutins, suberin). These specialised metabolites, deriving from secondary processes, can be used as flavours, aromas, nutraceuticals and drugs.

One of the first biochemical characterisation studies on plant P450s was conducted on avocado (*Persea americana*) fruits in 1989 by O'Keefe and Leto. Active P450 enzymes, named ARP1 and ARP2, belonging to the CYP71 family, were successfully obtained, through solubilisation of the microsomal fractions extracted from ripened avocado tissues (102).

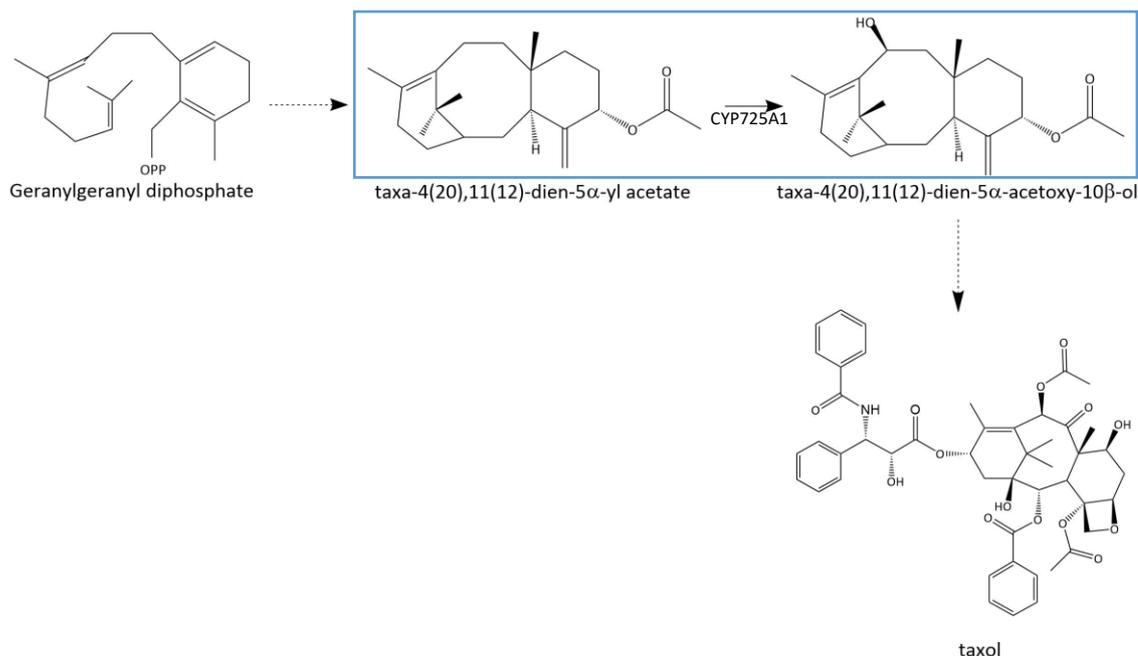
The 2015 Nobel Prize in Physiology or Medicine was awarded to scientist Tu Youyou for her discovery, in the early '70s, of the antimalarial properties of artemisinin, a

compound naturally produced in Artemisia plants (*Artemisia annua*) (103). The World Health Organization recognised this compound “as the first-line treatment for uncomplicated *Plasmodium falciparum* malaria”, in combined therapies. The very low amounts of artemisinin or precursors extractable from plants, alongside the economical unfeasibility of chemical synthesis, led to the long quest of researchers around the world towards the discovery of alternatives for the efficient production of the compound. Among the approaches used, are breeding programmes to obtain high-yielding hybrid plants and metabolic pathway engineering into different hosts for enhanced production. An example for the first type of approach is that developed by Prof. Graham’s group at the Centre for Novel Agricultural Products (CNAP), with the Artemisia Project, funded by the Bill & Melinda Gates Foundation (104). Genetic analyses and field trials led to the identification of key traits affecting the yields in artemisinic acid (which can be subsequently chemically converted to artemisinin); two hybrid Artemisia varieties have been produced and are now available on the market (105). On the other side, Paddon *et al.* manipulated and optimised the pathway leading to artemisinic acid (106, 107), by creating a synthetic yeast strain able to express, in parallel, enzymes that are directly involved in the pathway, as well as secondary factors, required to enhance the efficiency of the process. In particular, the authors, engineered the co-expression, with the CYP, of the partner reductase, two hydrogenases and  $\text{cyt}_{b5}$ . Competing enzymes were also inhibited. In this way, the authors succeeded in the production of up to 25 mg of artemisinic acid per liter of culture (107). This work was based on Ro and co-workers’ discovery that a P450 enzyme was a key player in the biosynthesis of artemisinic acid in Artemisia plants: CYP71AV1 is, in fact, responsible of three oxidation steps of the precursor amorphadiene to artemisinic acid (108, 109) (figure 1.15).



**Figure 1.15:** Reaction scheme illustrating the multiple steps catalysed by CYP71AV1, supported by CPR and  $\text{cyt}_{b5}$ , in the synthesis of artemisinin. (I): Amorpha-4,11-diene is hydroxylated by CYP71AV1 to artemisinic alcohol; (II): Artemisinic alcohol is oxidised by CYP71AV1 and artemisinic alcohol dehydrogenase 1 to artemisinic aldehyde; (III): Artemisinic aldehyde is converted by CYP71AV1 and artemisinic aldehyde dehydrogenase 1 to artemisinic acid, precursor of artemisinin.

Another plant metabolite with great pharmaceutical value is taxol, produced by yew trees (*Taxus* species) which is the active compound of Paclitaxel, a chemotherapeutic listed in the World Health Organisation “Model lists of essential medicines” (110). Schoendorf *et al.*, through differential display real-time PCR (DD-RT-PCR), identified a set of transcripts related to taxol biosynthesis. Within the set, were 13 P450 sequences, which showed a high sequence identity (between 52 and 83 %) with three families of CYPs (CYP83, CYP88, CYP90), some members of which have been already confirmed as active towards the synthesis of terpenoids (111). Schoendorf’s study demonstrated that one of the CYPs, taxane  $10\beta$ -hydroxylase (CYP725A1), expressed in yeasts fed with taxol precursors, was able to catalyse the conversion of taxadien- $5\alpha$ -yl acetate to taxadien- $5\alpha$ -acetoxy- $10\beta$ -ol (figure 1.16) (112).

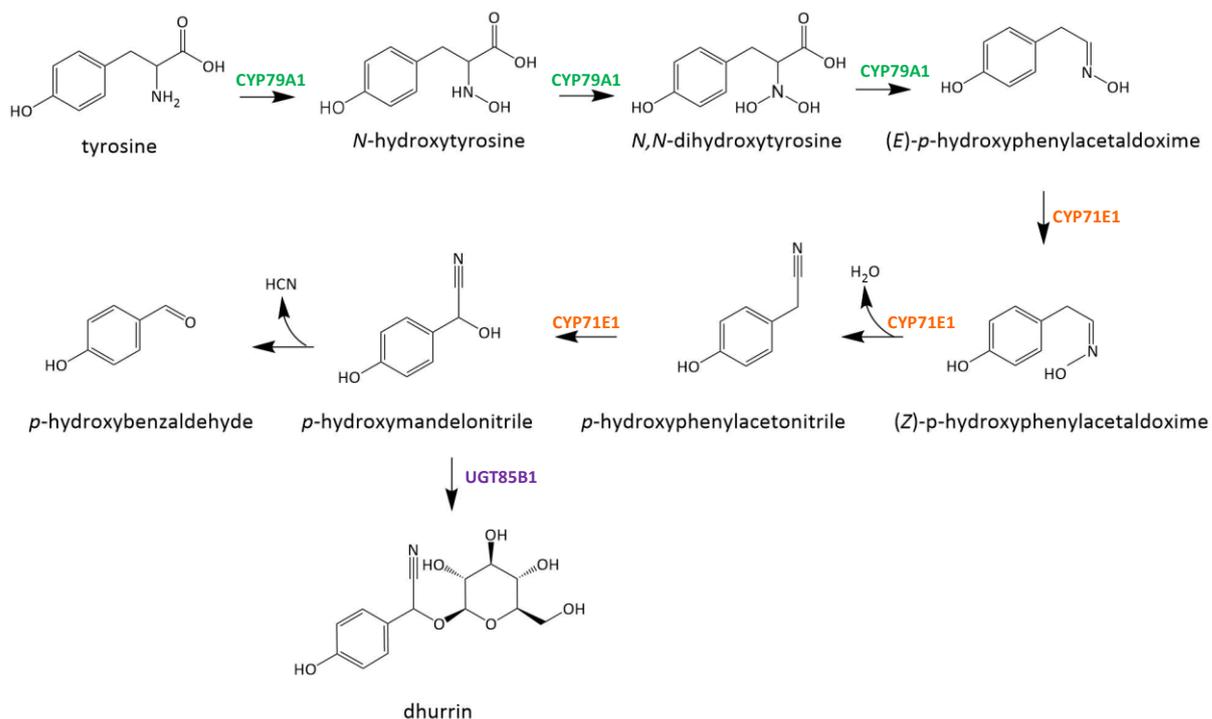


**Figure 1.16:** Abbreviated scheme of the taxol biosynthetic pathway from the geranylgeranyl diphosphate precursor. Highlighted in blue is the reaction catalysed by CYP725A1. The dashed arrows indicate missing intermediates not reported in the figure.

Later studies confirmed that of the 19 reactions leading from the geranylgeranyldiphosphate (GPP) precursor to taxol, half of the reactions are catalysed by P450 hydroxylases (113-115).

In an interesting proof-of-concept work, Lassen *et al.* demonstrated that photoautotrophic cyanobacteria can be employed as viable alternatives to the conventionally-used *E. coli* for gene expression. Additionally, PsaM, a section of photosystem I (PSI) in cyanobacteria, was exploited as a tag, in order to co-localise the target protein with the host's PSI in the thylakoids. The aim of this approach was to implement the electron flux required for P450 catalysis, by exploiting the photosynthetic system of the host cyanobacterium, complemented with ferredoxin, as also shown (and proven effective) in a previous study by Jensen *et al.* (116). In particular, Lassen and co-workers transformed the fusion of CYP79A1, one of the enzymes involved in the synthesis of cyanogenic glucosides in great millet (*Sorghum bicolor*), with PsaM into *Synechococcus* sp. PCC 7002. The CYP79A1-PsaM fusion was successfully expressed in the thylakoids; both *in vitro* and *in vivo* assays demonstrated that the complex was able to convert the substrate of CYP79A1,

tyrosine, into the expected oxime, exploiting energy from light (117). Subsequently, the whole dhurrin metabolon from *S. bicolor* (constituted by two CYPs, CYP79A1, CYP71E1 and a glucosyl transferase, UGT85B1 (118-120), see figure 1.17), was engineered into cyanobacteria (*Synechocystis sp.* PCC 6803).



**Figure 1.17:** Scheme representing the series of reactions carried out by the ‘dhurrin metabolon’, the dhurrin biosynthetic pathway in *S. bicolor*. CYP79A1 in three consecutive steps converts tyrosine into *p*-hydroxyphenylacetaldoxime. Subsequently, CYP71E1 catalyses the formation of *p*-hydroxymandelonitrile, which can degrade with release of cyanide. Alternatively, UGT85B1 can conjugate *p*-hydroxymandelonitrile with sugar, leading to dhurrin.

This synthetic engineering approach enabled the expression of the two CYPs as part of the cyanobacterial membrane, the expression of the soluble UGT, as well as the light-driven production of the expected final product, the cyanogenic glucoside dhurrin. This was achieved independently from the prohibitively-expensive NADPH and from the P450-reductase partner, which was efficiently replaced by the host’s photosynthetic system and by the supplied ferredoxin as the intermediate electron carrier (121). As demonstrated in these studies, cyanobacteria can be valuable hosts for expression, with the advantage of working as powerhouses to feed light-derived

energy to P450-mediated activities. This approach could be particularly promising for the synthesis of high-value chemicals.

The knowledge acquired on plant CYP enzymes can be turned into further societal benefit by exploiting those enzymes involved in the synthesis of defence compounds (cyanogenic glucosides) or in the breakdown of xenobiotics (herbicides, pesticides, or pollutants) for the development of resistant crops or for the improvement of the nutritional value of staple crops.

A transgenic *Arabidopsis* line was produced, which, by expressing the above-mentioned dhurrin metabolon (with the two CYPs and the UGT), displayed resistance against target herbivores (122). From another point of view, these studies can be the basis to solve another problem: food security. According to the US Department of Agriculture, cassava (*Manihot esculenta*) is a very important staple crop, ranking fourth after rice, maize and wheat (123). This plant is characterised by the ability to grow well in poor soils, it presents a very high tolerance to drought and the tubers can be left unharvested for a long time. This plant is mainly produced in developing countries, such as West Africa, tropical South America and South-East Asia, where it is used for its leaves (rich in vitamin A and B) and tuberous roots, as sources of carbohydrates. In addition, cassava, like great millet, produces high levels of poisonous hydrogen cyanide, released from cyanogenic glucosides (up to 1.5 g / kg), that are synthesised when the plant tissues are damaged by herbivores, but also during food preparation. All the parts of the cassava tuber have to be appropriately treated, to avoid poisoning, and, as a result, most of the nutrients are lost in the process. This factor is particularly negative in areas hit by food scarcity. Through sequence homology searches with the dhurrin pathway enzymes, it has been discovered that three CYPs, namely CYP71E7, CYP79D1 and CYP79D2, are involved in the synthesis of the two cyanogenic glucosides in cassava, linamarin and lotaustralin, from L-valine and L-isoleucine, respectively (124, 125). The identification of these enzymatic actors in the synthesis of cyanogenic glucosides constitutes an important starting point for the development of transgenic acyanogenic cassava plants, with decreased toxicity and higher nutritional value.

Exactly in the same way as in the human liver, cytochromes P450 in plants and bacteria are also able to metabolise xenobiotic compounds, such as pollutants and herbicides. These functions can be of high biotechnological relevance, as they can guide us towards new bioremediation strategies. An overview on this particular application of plant CYPs can be found in chapter 5.

## 1.7 Fungal Cytochromes P450

Fungal cytochrome P450 enzymes are involved in the biosynthesis of primary and secondary metabolites, with interesting biotechnological applications, and also in the degradation of xenobiotics. Considering that there are more than 2,500 fungal species, and also taking into consideration the high sequence variation generally displayed by CYPs across all domains of life, as well as Prof. Nelson's classification system (16), it is of no surprise that the overall number of fungal CYP families is high, with more than 330 annotated so far (126).

Fungal CYPs generally display a class II arrangement (see figure 1.5), with a FAD/FMN-binding CPR, transferring electrons to the heme from NADPH. Ichinose and Wariishi reported the unusual case of CYP5150A2 from *Phanerochaete chrysosporium*, which is able to catalyse the hydroxylation of diclofenac by receiving electrons from CPR. It was observed that the catalytic reaction was more efficient when cyt<sub>b5</sub> was added to the reaction mix, as well as when an alternative cyt<sub>b5</sub> reductase/cyt<sub>b5</sub> redox system was present, in place of the CPR (127). In addition, two peculiar CYP constructs have been found in fungi: P450<sub>nor</sub> (CYP55A4) and P450<sub>foxy</sub> (CYP505), from *Fusarium oxysporum*, both self-sufficient enzymes. In particular, P450<sub>nor</sub> is a single unit CYP, which catalyses the degradation of nitric oxide, conferring the denitrification capability to the organism. This activity does not require the presence of a redox partner, as the electrons are transferred directly from NAD(P)H to the heme, due to the composition and the conformation of the heme pocket. In the case of P450<sub>foxy</sub>, the P450 heme is directly fused to the redox partner. This membrane-bound fusion

enzyme shows a high degree of similarity, in terms of sequence and activity, with P450 BM-3 and functions as a fatty acid hydroxylase (128).

Additional characterised activities for fungal CYPs included those involved in the biosynthesis of sterols, mycotoxins and gibberellins and the catabolism of pollutants and antifungal compounds (126, 129, 130).

## **1.8 Challenges of using CYPs for biotechnological applications**

Since the first description in the early '60s, an astonishing amount of literature has been published so far on the cytochromes P450. A plethora of characterisation studies on P450s from different organisms have underlined the importance of CYPs, in the synthesis, metabolism or catabolism of a wide array of substrates, many of which have been discovered having valuable biotechnological applications. However, due to the complexity of the multicomponent P450 system, as well as the high occurrence of unproductive uncoupling reactions, the difficulty of producing recombinant stable activities and the generally poor product yields, the exploitation of this class of enzymes has been constrained. Besides the optimisation of the expression, purification, activity assay conditions and the manipulation of the P450 heme sequence, engineering need to be directed towards the enhancement of the electron transport system, through modifications at the redox partner level. In addition, optimal heterologous systems should be created, with "mix & match" combinations of P450 heme, redox partner and host. Furthermore, a significant amount of work has also to be carried out for the optimisation of the catalysis at both the biochemical and engineering levels (131).

## Aim of the project

This project aims to develop genetic constructs that would enable the stable and soluble expression of cytochromes P450 from plants. It is envisaged that these constructs could be utilised as drop-in vectors for the study of plant P450s with unknown function. Using the well-known CYP73A5 from *Arabidopsis* as a reference for validation, a particular focus is given to uncharacterised enzymes that could be potentially involved in the metabolism of xenobiotics. The constructs will include a plant reductase partner, which will be fused to the CYP domain, in order to obtain a redox self-sufficient fusion protein. Bacteria (*Escherichia coli*) will be the expression hosts of choice, due to their relative simplicity. In addition, yeasts (*Saccharomyces cerevisiae*) will also be employed, for comparative purposes. Finally, the obtained construct will be validated, by testing the activity of the inserted CYP73A5.

A section of the work is dedicated to the characterisation of the potential role of CYP81D11 in the detoxification of the explosive 2,4,6-Trinitrotoluene (TNT). Previous gene expression studies on plant tissues after exposure to TNT revealed that CYP81D11 was the most upregulated CYP, making it an interesting target for further study. Transgenic *Arabidopsis* plants and tobacco leaf discs will be employed for different types of assays. In particular, *Arabidopsis thaliana* plants overexpressing CYP81D11 will be incubated on liquid and agar media containing TNT to observe the presence of any activity, enhanced uptake or increased tolerance towards this compound, in comparison to wildtype and knockdown *Arabidopsis* plants. Furthermore, leaf discs from *Nicotiana benthamiana* leaves will be incubated with TNT to verify possible differences in the TNT metabolic pattern within the plant tissues, in comparison with controls.

## Chapter 2: General Materials and Methods

### 2.1 Chemical reagents

Reagents and consumables were purchased from: VWR (USA), Sigma-Aldrich (Poole, UK), Fisher Scientific (Loughborough, UK), Melford (Ipswich, UK), Formedium (Norfolk, UK), Duchefa Biochemie (Netherlands), New England Biolabs (USA), GE Healthcare (London, UK), Promega (Southampton, UK), Qiagen (West Sussex, UK), Expedeon (Swavesey, UK), Clontech Laboratories (USA), Invitrogen (Paisley, UK).

Pure water for the preparation of the media, the solutions and the assays was obtained with an Elga PureLab Ultra water polisher, from Elga Labwater (High Wycombe, UK).

### 2.2 Organisms

Bacterial strains: *Escherichia coli* strains DH5 $\alpha$ , Rosetta 2 (both from the Bruce group stocks), XL1-Blue, BL21(DE3) and Arctic express (from Agilent Technologies) were used for cloning steps and protein expression (table 2.1). *Agrobacterium tumefaciens* strain GV3101, obtained from the Bruce group stocks, was used for tobacco tissues infiltrations.

Yeast strain: *Saccharomyces cerevisiae* WAT11-modified (166), from the Bruce group stocks, was used for endogenous co-expression of ATR1 reductase.

Plants: *Arabidopsis thaliana* ecotype Columbia 0 (Col0) CYP81D11-modified lines (overexpressing and knock-down) for activity/phenotyping assays were obtained from Prof. J. Napier, Rothamsted Research; *Nicotiana benthamiana* was used for transient expression and activity assays, seeds were received from Luisa (Simon McQueen-Mason's group, York).

**Table 2.1:** Specifications of the bacterial strains used in the present work

Name	Genotype	Application	Endogenous Resistance and features
<b>DH5<math>\alpha</math></b> (Invitrogen™)	F <sup>-</sup> $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ ( <i>lacZYA-argF</i> ) U169 <i>recA1 endA1</i> <i>hsdR17</i> (rK <sup>-</sup> , mK <sup>+</sup> ) <i>phoA supE44</i> $\lambda$ - <i>thi-1</i> <i>gyrA96 relA1</i>	Cloning	None, High insert stability, high plasmid yield, blue/white screening of recombinant clones
<b>XL1-Blue</b>	<i>recA1 endA1</i> <i>gyrA96 thi-1 hsdR17</i> <i>supE44 relA1 lac</i> [F' <i>proAB lac</i> <sup>q</sup> $\Delta$ M15 Tn10 (Tet <sup>R</sup> )]	Cloning	Tetracycline, Blue/white color screening of recombinant clones
<b>Rosetta2</b>	F <sup>-</sup> <i>ompT hsdS</i> <sub>B(rB<sup>-</sup> mB<sup>-</sup>)</sub> <i>gal dcm</i> pRARE2 (Cam <sup>R</sup> )	Protein expression	Chloramphenicol, increased expression of eukaryotic proteins with rare codons
<b>ArcticExpress(DE3)</b>	B F <sup>-</sup> <i>ompT hsdS</i> <sub>B(rB<sup>-</sup> mB<sup>-</sup>)</sub> <i>dcm</i> <sup>+</sup> Tet <sup>R</sup> <i>gal</i> $\lambda$ (DE3) <i>endA</i> The[ <i>cpn10 cpn60</i> (Gent <sup>R</sup> )	Protein expression	Gentamycin, constitutive expression of chaperonins

## 2.3 Media

### Media for bacterial growth

Cultures of *E. coli* were firstly grown in nutrient media, such as Lysogeny broth (LB) or Super Optimal broth with Catabolic repressor (SOC) or Y medium and later sub-cultured in Auto Induction (AI) medium or M9 minimal medium for gene expression.

Plasmid transformation and competent cell growth were performed in SOC medium. All the antibiotics for selection were prepared at a working concentration of 50-100 µg/ml and filter-sterilised prior to use.

The LB medium consisted of 5 g/l yeast extract, 10 g/l tryptone and 10 g/l sodium chloride. 1% w/v of agar was added for the preparation of LB agar plates.

The SOC medium was obtained by adding 10 ml/l of filter-sterilized 2 M glucose to SOB medium, just prior to use. SOB medium composition: 20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l sodium chloride. The solution was autoclaved and 10 ml of filter-sterilised 1 M magnesium chloride and 10 ml of filter-sterilised 1 M magnesium sulphate were finally added to obtain the complete SOB medium.

The Y medium consisted of 5 g/l yeast extract, 20 g/l of tryptone, 5 g/l magnesium sulphate, pH 7.6.

The AI medium was a complex medium containing 10 g/l tryptone, 5 g/l yeast extract, 20 ml/l 50X 5052 solution (250 g/l glycerol, 25 g/l glucose, 100 g/l  $\alpha$ -lactose), 50 ml/l NPS solution (1.8 g of sodium sulphate, 8.5 g  $\text{KH}_2\text{PO}_4$ , 8.85 g  $\text{Na}_2\text{HPO}_4$ , 6.7 g ammonium chloride dissolved in 45 ml of  $\text{H}_2\text{O}$ ), 1 ml/l 1 M magnesium sulphate, 1 ml 1000x trace metals (0.1 M  $\text{FeCl}_3$ , 1 M  $\text{CaCl}_2$ , 1 M  $\text{MnCl}_2$ , 1 M  $\text{ZnSO}_4$ , 0.2 M  $\text{CoCl}_2$ , 0.1 M  $\text{CuCl}_2$ , 0.2 M  $\text{NiCl}_2$ , 0.1 M  $\text{Na}_2\text{MoO}_4$ , 0.1 M  $\text{Na}_2\text{SeO}_3$ , 0.1 M  $\text{H}_3\text{BO}_3$  in water).

The M9 medium (5X concentrated) was composed of 33.9 g/l of  $\text{Na}_2\text{HPO}_4$ , 15 g/l  $\text{KH}_2\text{PO}_4$ , 2.5 g/l sodium chloride, 5 g/l ammonium chloride. The medium, diluted down to 1X strength with Elga water, was complemented with filter-sterilised 1 M glucose (10 ml/l), 1 M  $\text{CaCl}_2$  (100 µl/l) and 1 M magnesium sulphate (2 ml/l) right prior to use.

### **Media for plant growth and assays**

Murashige & Skoog (1962)(MS, used at ½ strength) was prepared with 2.15 g/l MS powder and 20 mM sucrose. The pH value was corrected to 5.7 with sodium hydroxide prior to autoclave sterilisation.

## 2.4 Molecular Biology techniques

### 2.4.1 Polymerase Chain Reaction (PCR) & colony PCR

All the oligonucleotides used as primers for the PCR reactions were designed using the OligoAnalyzer online tool (<https://www.idtdna.com/calc/analyzer>), provided by Integrated DNA Technologies (IDT) and synthesized by either IDT or Sigma. The chain reactions were performed on Bio-Rad DNA Engine thermal cyclers, using *Pfu*, Phusion, Q5 and GoTaq polymerases, following the instructions and protocols provided by the manufacturer. The colony PCR was performed using single bacterial colonies (previously transformed with a plasmid and grown on antibiotic-containing medium), instead of pure DNA, as templates for the amplification cycles.

### 2.4.2 DNA restriction digestion

The destination vectors were linearised by single or double digestion with appropriate endonuclease enzymes: *NcoI*, *XhoI*, *AvrII*, *SacI*, *BamHI*, *BseRI* (from NEB), mainly. As indicated by the product's guidelines, all the restriction reactions, with 1 U of enzyme /  $\mu\text{g}$  of DNA in the provided CutSmart buffer, were performed at 37 °C for 30 min – 2 h.

### 2.4.3 Dephosphorylation of linearised vector

To avoid the re-annealing of the DNA ends generated by the digestion, the linearised vectors were treated with FastAP alkaline phosphatase (ThermoScientific), which allowed the removal of the overhanging phosphate groups. As suggested by the manufacturer, the dephosphorylation reactions were carried out at 37 °C for 10 min and the enzyme was subsequently inactivated by heating at 75 °C for 5 min.

### 2.4.4 Agarose gel electrophoresis

The size of the PCR amplicons and the effective linearisation of the vector were analysed via gel electrophoresis. The samples, mixed with the stain, were loaded on a 1.2 % w/v agarose gel, prepared with TAE buffer (for 50X strength: 242 g/l of Tris

base, 57.1 ml/l of glacial acetic acid, 100 ml/l of 0.5 M EDTA) and supplemented with ethidium bromide (2  $\mu$ l/100 ml buffer) to visualise the DNA. A BioRad PowerPac 3000 was employed to generate the current, generally 100-140 V, for the migration of the analytes, charged negatively, towards the cathode. The sizes of the analytes were compared to a DNA ladder, generally a 1 kb DNA Ladder (from NEB, catalogue number N3232 or Promega, catalogue number G5771, with markers sized 0.5-10 Kilobases). A gel documentation imaging system, equipped with a UV chamber and a camera lens, was used to visualise the DNA and record a picture of the gel.

#### **2.4.5 DNA isolation**

The PCR amplicons and the digested plasmids were isolated using the Wizard<sup>®</sup> SV Gel and PCR Clean-Up System (Promega), following the provided protocol. Briefly, the DNA samples were mixed with the membrane binding solution (1:1 ratio) and loaded onto the Minicolumn. After a 1-min incubation, centrifugation at 16,000 g for 1 min allowed the separation of the linear DNA present in the sample from the buffer and residual chemicals/enzymes from previous steps. The immobilised DNA was subsequently washed with an ethanol-based wash buffer and the final, clean, DNA was then eluted with nuclease-free water.

For the isolation of plasmids, a QIAprep Spin Miniprep kit (Qiagen) was employed, following the instructions from the manufacturer. The procedure, similar to that used for the purification of linear DNA, was preceded by a lysis step, where cells transformed with the plasmid were disrupted chemically to allow the release, among other cellular contents, of the plasmid of interest. After a centrifugation step, at 17,900 g for 1 min, the plasmid was loaded onto a spin column with a selective membrane and separated from the cell debris via centrifugation, at 17,900 g for another min. Two ethanol-based washing steps followed (at the same centrifugation speed and time), with a final elution of the plasmids with pre-heated nuclease-free water.

In both cases the isolated DNA was kept on ice for immediate use or stored at -20 °C.

### **2.4.6 Cloning: ligation, InFusion**

Different procedures were employed to introduce the insert of interest in the destination plasmid. The traditional method consisted of ligating the compatible ends of the insert with those of a linearised vector using a T4 DNA ligase (NEB), which catalyses the formation of phosphodiester bonds. According to the protocol provided, the reaction mixtures, with 1  $\mu$ l of enzyme, corresponding to 400 NEB Units, and different vector: insert ratios (1:1, 3:1, 1:3), were incubated for up to 2 h at room temperature and then the enzyme was inactivated at 65 °C. InFusion cloning (Clontech) is an alternative method for cloning inserts into vectors. Using the InFusion kit, the desired insert was PCR-amplified for the introduction of 15 basepair-long 3' and 5'-flanking adapters homologous to the ends generated by enzymatic digestion of the destination vector. The linearised vector and the "adapted" insert were mixed (insert:vector ratio of 2:1) with 2  $\mu$ l of the InFusion enzyme premix and incubated at 50 °C for 15 min. At the end of both cloning reactions (ligation and InFusion) the mixtures were straightaway transformed into competent DH5 $\alpha$  or XL1-Blue *E. coli* cells.

### **2.4.7 Preparation of competent cells**

A chemical treatment was used for the preparation of competent *E. coli* cells, in order to allow the efficient transformation of exogenous DNA of interest.

A single colony 5 ml inoculum in Y medium was firstly incubated at 37 °C. After allowing a first round of cell replication, the starter culture was seeded into 250 ml of SOB medium in a 2 l Erlenmeyer flask, to allow a good overall aeration of the culture. The flask was left on the orbital shaker (set at 200 rpm) at 18 °C until the optical density of the culture (OD<sub>600</sub>) reached a value of 0.6. The cells were harvested by centrifugation at 5,000 g for 15 min at 4 °C. The cell pellets were firstly treated with a medium composed of 30 mM potassium acetate, 100 mM rubidium chloride, 10 mM calcium chloride, 50 mM manganese chloride, 15 % v/v glycerol, pH 5.8. After a 15-min centrifugation at 5,000 g at 4 °C the cells were re-suspended in 10 mM MOPS, 75 mM calcium chloride, 10 mM rubidium chloride, 15% v/v glycerol. The

resulting cell suspension was aliquoted, snap frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

## **2.4.8 DNA transformation**

### **2.4.7.1 Heat shock**

The *E. coli* competent cells, briefly let to thaw on ice, were mixed with  $1\ \mu\text{l}$  (approx. 50 ng) of plasmid DNA. After 10 to 30 min of further incubation on ice, the mix was heated for 25 (for Arctic express) or 60 s (for all the other strains) on the water bath, set at  $42^{\circ}\text{C}$ . The mix was then placed back on ice for two min and the cells were recovered with pre-warmed nutrient medium (SOC or LB). After a 1- h incubation at  $37^{\circ}\text{C}$  on the orbital shaker (set at a speed of 180 rpm), the transformed cells were spread onto LB -agar plates. The antibiotic in the medium was matched to the presence of the antibiotic resistance gene encoded on the transformed vector, to allow the selection of the transformed cells.

### **2.4.7.2 Electroporation**

A  $50\ \mu\text{l}$  aliquot of competent *A. tumefaciens* (strain GV3101) was first mixed with  $1\ \mu\text{l}$  of vector DNA. After a brief incubation on ice, the cells were transferred into a 2 mm micropulsar cuvette and electroporated using a BioRad Gene Pulser II system. Electroporation setting: 2.5 kV,  $400\ \Omega$  resistance,  $25\ \mu\text{F}$  capacitance. The cells were recovered with 1 ml of LB medium, incubated for 3 h on the orbital shaker (200 rpm) at  $30^{\circ}\text{C}$  and spread onto LB agar plates, with the appropriate antibiotic for resistance selection.

## **2.4.9. DNA sequencing**

The purified DNA samples at a concentration of 80-100 ng/ $\mu\text{l}$  were sequenced either via the departmental Genomics Technology Facility or through the external GATC Biotech AG (Germany) sequencing provider.

## 2.5 Biochemistry techniques

### 2.5.1 Protein heterologous expression

#### 2.5.1.1 In *E. coli* Rosetta2, Arctic express, BL21(DE3)

Many different conditions were tested for the soluble expression of the target proteins; specific conditions are supplied in the relevant Chapters. In general, expression cultures were prepared in M9 minimal medium, with glucose as the carbon source. At an OD<sub>600</sub> value between 0.6-0.8, 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG), 0.5 mM aminolevulinic acid and 0.5 mM FeCl<sub>3</sub> were added, to induce protein expression and as heme precursors, respectively. The cultures were then incubated on the orbital shaker (180 rpm) for 16-20 h at 16 °C.

#### 2.5.1.2 In *A. tumefaciens*

Single colonies grown on the selection plates were inoculated in 10-20 ml of LB medium, containing the appropriate antibiotic. The cultures were left to grow overnight at 30 °C on the orbital shaker (180 rpm).

### 2.5.2 Cell harvest & lysis

At the end of the expression period, the cells were harvested via centrifugation at low speed (5,000 g) at 4 °C on a Sorvall centrifuge. The cells were then resuspended in lysis buffer, generally a sodium or potassium phosphate buffer at pH 7.4-8, supplemented with 200 μM of phenylmethylsulfonyl fluoride (PMSF). The suspensions were stirred for 30 min in the cold room and then subjected to sonication bursts (using a Misonix S-4000 sonicator) at 70 % amplitude, 3 s on and 7 s off, for a total sonicating time of 4 min, to disrupt the cells and allow the release of the proteins. The soluble components were separated from the cell debris via centrifugation at 34,000 g at 4 °C for 15 min.

### 2.5.3 Spectrophotometric evaluation of protein expression

The presence of the target P450 proteins in the clarified lysates or at the end of the purification process could be verified using a UV-Vis spectrophotometer, by scanning within a wavelength window between 300 and 600 nm. If a cytochrome P450 is present in the mixture, an absorbance peak at 420 nm (related to the P450 enzyme in the oxidised resting state) should be clearly observed. To further verify the correct folding of the expressed protein, 5 mg of sodium dithionite were added in the sample cuvette to reduce the protein. The sample was then “bubbled” with carbon monoxide (1 bubble/s for 1 min, approximately). The final CO-bound state, indicative of a correctly folded protein, should present an absorption maximum at 450 nm (4, 133).

### 2.5.4 SDS-PAGE verification

Samples containing the protein of interest were mixed with Laemmli loading buffer (134), 3 parts protein: 1 part buffer, composed of 15 % (v/v) distilled water, 25 % (v/v) 1M Tris-HCl, pH 6.8, 40 % (v/v) glycerol, 10 % (w/v) sodium dodecyl sulfate (SDS), 0.05% (w/v) bromophenol blue and 20 % (v/v)  $\beta$ -mercaptoethanol. After a 5-min incubation at 90 °C, the samples were loaded on a 12 % polyacrylamide gel. The gels were casted right prior to loading; the separating section of the gel was composed of 1.68 ml water, 1.25 ml 1.5 M Tris-HCl pH 8.8, 50  $\mu$ l 10 % (w/v) SDS, 2 ml Protogel 37.5:1 acrylamide to bisacrylamide stabilised solution (National Diagnostic), 37.5  $\mu$ l 10 % (w/v) ammonium persulphate (APS), 3.75  $\mu$ l *N*-Tetramethylethylenediamine (TEMED). The stacking section of the gel, comprising the loading wells, was instead made of 1.22 ml water, 0.5 ml 0.5 M Tris-HCl pH 6.8, 20  $\mu$ l 10 % w/v SDS, 0.26 ml Protogel solution, 15  $\mu$ l APS, 3  $\mu$ l TEMED. The electrophoresis was carried out at a voltage of 200 V in a BioRad miniprotean 3 cell, filled with running buffer, composed of 3 g/l Tris, 14.4 g/l glycine, 1 g/l SDS. At the end of the run, the separated analytes present in the sample were visualised by incubating the gel for 30 min on the rocker at room temperature with 20 ml of Expedeon Instant blue protein stain. The size of the analytes was compared with molecular markers, mainly NEB PageRuler plus, which covered the 10-250 kDa range of sizes.

### **2.5.5 Western blot verification**

At the end of the electrophoretic run, the SDS acrylamide gel was positioned on a nitrocellulose membrane (BioRad, 0.45  $\mu\text{m}$ ) and the two layers were encased in six total layers (three per side) of thick filter paper pre-soaked in Towbin transfer buffer (25 mM Tris, 192 mM glycine, 20% v/v methanol). The resulting “sandwich” was positioned within an electrified circuit (BioRad Transblot SD SemiDry Transfer cell) set up at 25 V, allowing the transfer of the analytes from the gel onto the membrane. The effective blotting on the membrane was verified via Ponceau staining, incubating the membrane briefly with a few ml of Ponceau S solution (0.1% w/v Ponceau S in 5% v/v acetic acid). The membrane was incubated for 1 h in PBS buffer (137 mM NaCl, 2.5 mM KCl, 10 mM  $\text{Na}_2\text{HPO}_4$ , 1.5 mM  $\text{KH}_2\text{PO}_4$ , pH 7.4) containing 5% w/v of skimmed milk powder. After this blocking step, the membrane was washed three times with PBS buffer containing 0.5% of Tween20 and incubated for at least 2 h with PBS buffer solution containing 1% BSA and an AntiHis polyclonal antibody conjugate from rabbit (1/2,000 dilution). At the end of the incubation period the membrane was rinsed three times with the 0.5 % Tween20 solution and finally developed, revealing the His-tagged conjugated products, using a developing solution composed of 10 ml triethanolamine, 2 ml chloro-naphtol in methanol and 5  $\mu\text{l}$  of hydrogen peroxide. Authentic His-tagged proteins were also loaded on the SDS-PAGE gels, as positive controls.

### **2.5.6 Protein purification**

The proteins present in the clarified cell lysates were purified via affinity chromatography, exploiting the affinity of the polyHistidine tags for the Nickel-covered resin, either with a batch process or by employing automated pump systems (peristaltic pump or Akta purifier).

Prior to the purification the clarified cell lysates were filtered through a 0.22  $\mu\text{M}$  Millex sterile filter unit (Millipore).

### **Batch process**

Small scale purification trials were performed in batch systems, consisting of pre-equilibrated His-select Nickel affinity gel (0.5 ml / 10 ml of lysate) in a 50 ml Falcon tube. Briefly, the filtered lysate was incubated for more than 2 h (up to overnight) with the pre-equilibrated resin on the rocker. At the end of the incubation, the mix was centrifuged at 5,000 g for 10 min. The buffer was replaced (gently, using a pipette) with 15 resin volumes of fresh buffer, to remove the unbound proteins. After a brief mix and another centrifugation step (same conditions as above), the resin was washed again two more times. One resin volume of fresh buffer containing 5 mM imidazole was used to displace non-specifically bound proteins from the resin. After this fourth washing step, the proteins of interest were detached by incubation with one resin volume of buffer containing 500 mM imidazole. The mixture was incubated shaking gently for 1 h on the rocker and then centrifuged, in order to settle down the resin. The supernatant was collected and loaded onto a spin column (0.22  $\mu$ m in cellulose acetate, Agilent Tech), and centrifuged for 2 min at 16,000 g, to remove the remaining traces of resin. The protein solution was then dialysed twice against 5 l of potassium phosphate buffer pH 7 using Slide-A-Lyzer™ G2 Dialysis Cassettes (ThermoFisher Scientific). At the end of the process the protein solution was recovered, mixed 1:3 with 60 % (v/v) sterile glycerol, snap frozen in liquid nitrogen and stored at – 80 °C.

### **Continuous process**

The filtered lysate was loaded at a rate of 0.5 ml / min onto a pre-equilibrated HisTrap FF crude (GE Healthcare Life Sciences) 1 ml / 5 ml chromatographic column, connected either to a peristaltic pump or to an ÄKTApurifier 10 system. Elution was carried out with a gradient of 30-300 mM imidazole, with a 1 ml/min flow. The eluted fractions were checked spectrophotometrically and those presenting the higher 420 nm peaks (related to the CYP in the oxidised resting state) were pooled together, uploaded into Slide-A-Lyzer™ G2 Dialysis Cassettes and dialysed overnight in potassium phosphate buffer pH7, to minimize the imidazole concentration. The recovered protein was then concentrated using a centrifugal concentrator, generally a Vivaspin® 2 Centrifugal Concentrator. Purity and size of the protein in the collected

samples were determined using SDS polyacrylamide gel electrophoresis. The absorbance spectrum between 300 and 600 nm was measured to verify the presence of the characteristic peak at 420 nm.

### **2.5.7 Protein quantification**

The quantity of protein present in solution was evaluated either directly, inputting the 280 nm absorbance read with a NanoDrop-1000 spectrophotometer in the Lambert Beer's law, with the known extinction coefficient, or indirectly via Bradford assay. The latter was performed in a 96-wells format mixing 10  $\mu$ l of sample (diluted appropriately in order not to exceed the sensitivity of the assay) with 300  $\mu$ l of Coomassie dye. After a 10-s shaking, the absorbance at 595 nm, related to the dye-protein complex, was read using a Tecan Sunrise™ microplate absorbance reader. The values were compared to those of bovine serum albumin (BSA) standards of known concentrations, ranging from 0 to 2 mg/ml, to calculate the final concentration of the protein present in solution.

### **2.5.8 Activity assay**

#### **CYP73A5 activity towards cinnamic acid**

After 24 h of growth at 16 °C the cells were harvested via centrifugation and resuspended in 50 mM potassium phosphate buffer pH 7 (100 mg cells / ml of buffer). The suspension of whole cells was used to assess the ability of the construct to catalyse the hydroxylation of cinnamic acid, the natural substrate of CYP73A5, to *p*-coumaric acid. The reaction, carried out at 28 °C with shaking (300 rpm) was initiated by the addition of 200  $\mu$ M cinnamic acid and samples were collected over a period of 7 h. Reactions were terminated with methanol and the supernatant, after a centrifugation step to remove the cells, was analysed via HPLC with a Waters X-Bridge C18 column (250 x 4.6 mm, 5  $\mu$ M) heated to 30 °C, employing a gradient of water (A) and methanol with 0.1% acetic acid (B): 90 % A for the first 3 min, 90 to 55 % A in 1 min, 55 % A for 5 min and 90 % A for 5 min. The HPLC system used consisted of a Waters Alliance 2695 separation module and a Waters 2996 photodiode array detector.

## **Chapter 3: Soluble expression of Cytochromes P450**

### **3.1 Introduction: Bottlenecks in the heterologous expression of plant CYPs**

Eukaryotic cytochrome P450 enzymes, unlike their prokaryotic counterparts, are tightly bound to the membranes of the endoplasmic reticulum through an N-terminal hydrophobic domain. This means that when recombinantly expressed, eukaryotic cytochromes P450 are almost always insoluble. This insolubility issue has constituted a significant constraint in the process of discovering, and characterising, novel cytochrome P450 functionalities from eukaryotes. Generally, the native expression of plant CYP enzymes is very low, and tightly dependent on the developmental stage of the organism or on environmental inducers. Thus, many protein characterisation studies have been, and still continue to be, carried out by overexpressing genes of interest heterologously. Among the favourite hosts of expression are bacteria, such as *E. coli*, which are relatively simple and inexpensive systems.

Several approaches have been used so far to tackle the problematic expression of eukaryotic enzymes in bacterial hosts. These include nucleotide sequence optimisation re-coding, to overcome codon usage bias, and protein sequence modification, to remove hydrophobic domains or amend the structural arrangement. Additionally, expression conditions and purification processes have to be optimised, depending on the characteristics of the target protein, such as the presence of tags, isoelectric point, stability. The first studies reporting the successful expression of eukaryotic CYPs in a bacterial system were published in 1991. Li and Chan succeeded in obtaining high levels (up to 75%) of soluble and active rat hepatic cholesterol 7 alpha-hydroxylase cytochrome P450 (P450c7) in *E. coli*. This successful expression was achieved by deleting the amino acids in positions 2 to 24 from the starting methionine; amino acids identified as forming the hydrophobic anchor from the N-terminus (135). Barnes and co-workers (136) next obtained high levels of active

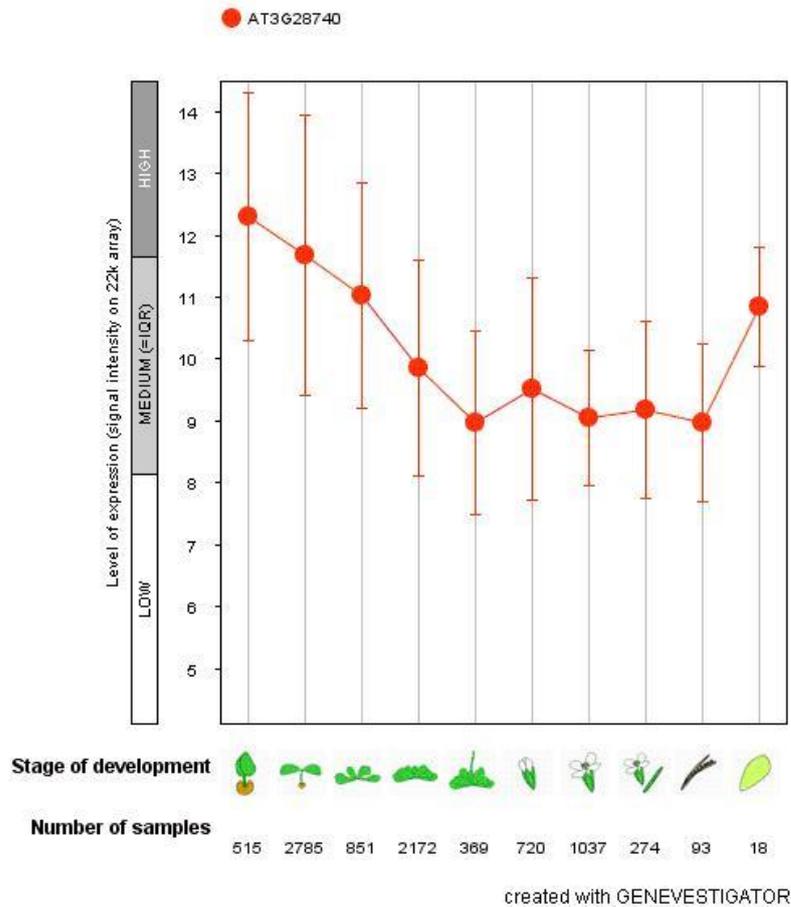
bovine CYP17 $\alpha$ -hydroxylase, a CYP that is involved in hormone biosynthetic pathways, using *E. coli* JM109 as the host for expression. Key to Barnes *et al.*'s success was the use of a strong, inducible promoter, as well as the modification of the N-terminus of the protein. Specifically, a tryptophan was replaced with an alanine and four other codons were subjected to silent mutations. The authors explained that these mutations minimised the formation of secondary structures in the mRNA and adhered to the codon preference of the host organism, significantly improving the efficiency of the host ribosomal translation (136, 137). Subsequently, other groups demonstrated that simply replacing the N-terminal hydrophobic anchor with the eight amino acids derived from the modified bovine CYP17 $\alpha$  (MALLLAVF) paved the way for expression and characterisation of many mammalian P450s such as human CYP1A2 (138), rabbit CYP2C3 (139) and human CYP2C10 (140). It is of note that in all these cases, the MALLLAVF peptide yielded high levels of CYPs, however, these products were all bound to the membranes in the host. Additional solubilisation steps with detergents were therefore needed to isolate the proteins. An alternative substitutive N-terminal peptide described in the literature was MAKKTSSKGKL, engineered for the first time by von Wachenfeldt *et al.* (141) and which allowed the expression, as well as functional and structural studies, of several eukaryotic CYPs such as rabbit CYP2C3 (141), human CYP2D6 (142), CYP2B6 (143), CYP2C8 (144), CYP2C9 (145), CYP2C5 (146, 147), CYP2R1 (148) and also Arabidopsis CYP74A1 (149), CYP98A3 (150) and burdock (*Medicago truncatula*) CYP93C20 (151).

Dr. Julia Schuckel, a previous member of the Bruce group, attempted the expression of a set of plant CYPs – CYP81D11, CYP81D8 from Arabidopsis and CYP71A12 from peppermint (*Mentha x piperita* L.) - either by deleting the N-terminal anchor (CYP81D11tr and CYP81D8tr) or by replacing it with the bovine peptide mentioned above (CYP71A12 PM2-2). Additionally, she performed several expression screens, employing a broad range of different hosts (*E. coli* BL21(DE3), Rosetta 2, Rosetta gami 2, DH5 $\alpha$  strains and *S. cerevisiae*, WAT21 and WAT11 modified strains) and media (TB, M9 and LB), but the levels of expression were not sufficiently high to allow western blot verification of the purified products (152). Following the published procedure by Haudenschild *et al.* (153), CYP71A12 PM2-2 was transformed in *E. coli*



To date, *AtCYP81D11* is an uncharacterised enzyme, produced at medium basal levels throughout the different growth stages and in different tissues of the *Arabidopsis* plant (see GeneVestigator analysis in figures 3.2 and 3.3).

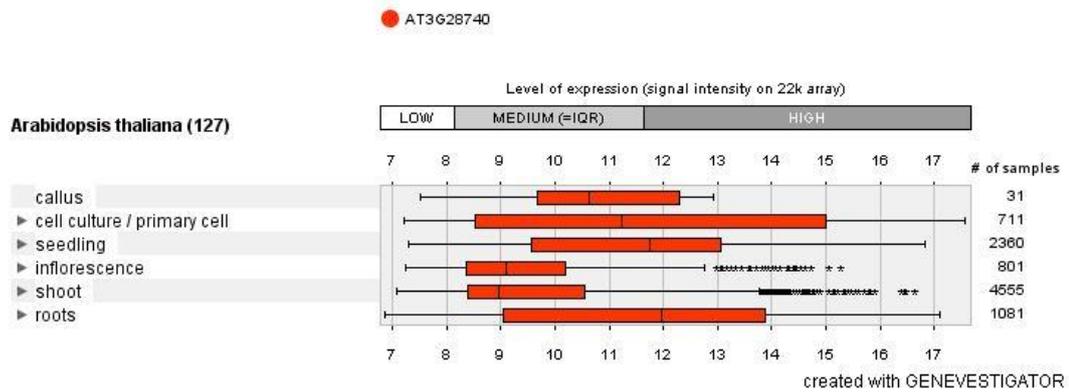
**Dataset:** 10 developmental stages (sample selection: AT-SAMPLES-0)  
 1 gene (gene selection: AT-GENES-0)



**Figure 3.2:** Expression pattern of CYP81D11 in *Arabidopsis* at different developmental stages. Source: <https://www.genevestigator.com/gv/>, 15/07/2014.

### Chapter 3: Soluble expression of CYPs

**Dataset:** 127 anatomical parts (sample selection: AT-SAMPLES-0)  
 1 gene (gene selection: AT-GENES-0)



**Figure 3.3:** Expression pattern of *AtCYP81D11* in different tissues. Source: <https://www.genevestigator.com/gv/>, 15/07/2014.

Many studies have reported how CYP81D11 expression levels were upregulated in plants subjected to different types of stresses, such as bacterial infections, wounding, exposure to toxic compounds, extreme temperatures, drought (159). Interestingly, as highlighted by Ekman *et al.* (160), as well as in Gandia-Herrero *et al.* (161, 162), the gene encoding CYP81D11 is the most upregulated, up to 25 fold higher levels, upon treatment of *Arabidopsis* plants with the explosive 2,4,6-trinitrotoluene (TNT). Other upregulated genes identified in Gandia-Herrero *et al.*'s study (162), such as OPRs, GSTs and UGTs, were discovered to be directly active towards TNT or its derivatives (163, 164). Additionally, oxidative derivatives of this xenobiotic compound have been described in the literature (165). These observations might thus lead to the possible involvement of CYP81D11 in the detoxification of TNT.

The *ZmCYP81A9* shares a 79 % sequence identity with CYP81A21 and CYP81A12, two enzymes from *Echinochloa phyllopogon*, a pervasive weed of rice. Both *EpCYPs* are able to catalyse the demethylation of bensulfuron-methyl, an acetolactate synthase-inhibiting herbicide (166). Another match (73% sequence identity) was found for rice (*O. sativa*) CYP81A6, which was reported as responsible for herbicide resistance in rice (167). An RNAseq study led by Liu *et al.* (168) showed that CYP81A9 is among the most upregulated genes upon treatment of maize plants with the herbicide

nicosulfuron, thus implying a potential role of this enzyme in the plants' defence mechanisms against xenobiotics. For this reason, *ZmCYP81A9* was chosen as an interesting target for potential agronomical purposes.

The aim of this section is to optimise the production, via protein engineering and multiple expression screenings, of the selected set of plant CYPs.

## 3.2 Materials and methods used for the soluble expression of plant CYPs

### 3.2.1 Sequence analysis, PCR and cloning

The native coding sequences for the target P450 enzymes were retrieved from the National Center for Biotechnology Information (NCBI) database (table 3.1).

**Table 3.1:** Gene ID and locus tags of the proteins used in this chapter

Protein	Role	Gene ID - locus tag
<b><i>AtCYP73A5</i></b>	Hydroxylation of cinnamic acid (phenylpropanoid pathway)	817599 - AT2G30490
<b><i>AtCYP81D11</i></b>	Xenobiotic metabolism (hypothesised)	822506 - AT3G28740
<b><i>ZmCYP81A9</i></b>	Xenobiotic metabolism (hypothesised)	103625869 - ZEAMMB73_713160
<b><i>AtCYP81D8</i></b>	Xenobiotic metabolism (hypothesised)	829891 - AT4G37370

The native plant nucleotide sequences were codon-optimised for expression in *E. coli* using the 'GeneOptimizer' online tool (<https://www.thermofisher.com/uk/en/home/life-science/cloning/gene-synthesis/geneart-gene-synthesis/geneoptimizer.html>) by ThermoScientific. The ExPASy 'protein translate' online tool ([web.expasy.org/translate/](http://web.expasy.org/translate/)) was employed to convert the nucleotide sequences into amino acid sequences. The obtained

sequences were then analysed using the TMHMM ([www.cbs.dtu.dk/services/TMHMM](http://www.cbs.dtu.dk/services/TMHMM), Centre for Biological Sequence Analysis, Denmark Technical University) and the DAS ([www.sbc.su.se/~miklos/DAS](http://www.sbc.su.se/~miklos/DAS), a joint project by the Biological Research Center - Hungarian Academy of Sciences and the Stockholm University) online transmembrane domain prediction tools, in order to identify the possible N-terminal membrane-spanning regions. These tools employed two different algorithms to determine protein domains with affinity to the endoplasmic reticulum (ER). Removal of the DNA sequences encoding these domains could decrease the affinity of the resultant proteins to the ER membrane and thus increase their overall solubility. The resulting optimised and truncated sequences of CYP81D11 (*CYP81D11tr*) and CYP81A9 (*CYP81A9tr*) were synthesised using the GeneArt™ gene synthesis service (Life Technologies). The *CYP73A5tr* sequence was already available in the Bruce group's stock.

A new expression construct, used in the work presented here, was designed specifically to express plant CYPs. This construct encoded the MAKKTSSKG peptide, hereafter named the S-tag, at the N-terminus of the target protein's reading frame. In order to introduce this peptide "signal" (which would presumably decrease the protein's affinity for the ER membrane), an intermediate cloning vector containing the peptide was synthesised by GeneArt™, called pMA-T. In addition to the N-terminal tag, pMA-T also presented, within the cloning site, other strings of aminoacids, to subsequently use as a linker to connect the CYP and the redox partner in the fusion constructs (Chapter 4).

The primers used to amplify the CYP sequences throughout the cloning process are summarised in table 3.2.

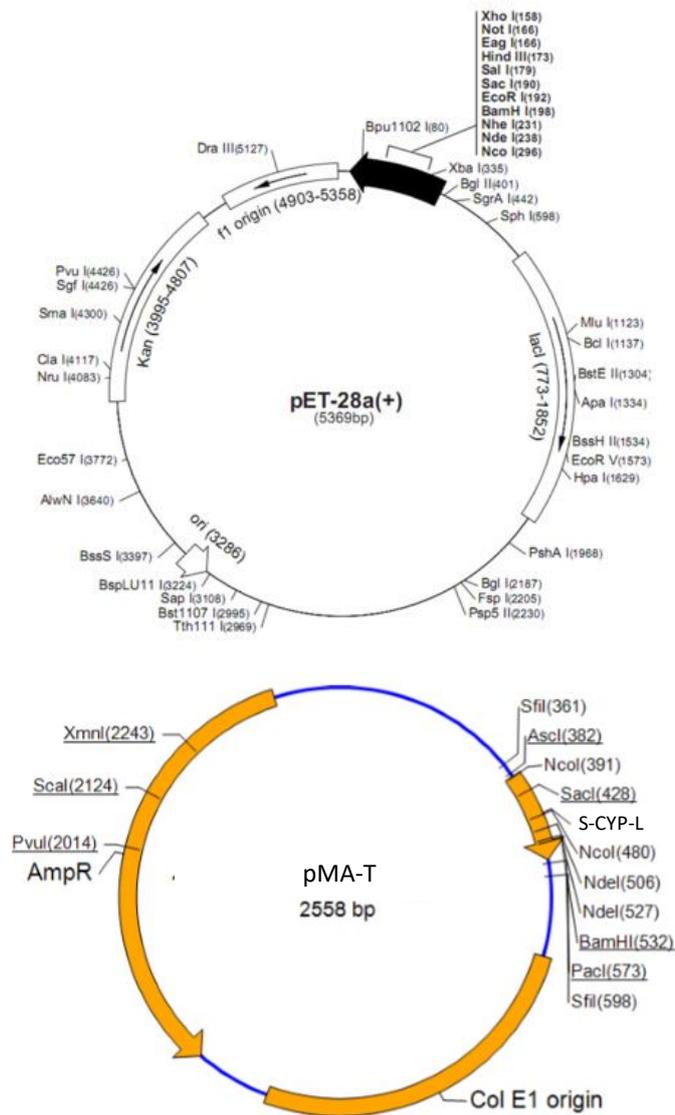
**Table 3.2:** Oligonucleotides used for the PCR amplification of the three CYP targets.

Protein name	Primers
<b>CYP73A5-1</b>	Forward - 5' AGCAAAGGTGAGCTCGCCCTGCGTGGTAAAAAAC 3'
	Reverse - 5' GCTGGTGCTACCGGATCCTGAGTTACGCGGTTTCAT 3'
<b>CYP73A5-2</b>	Forward - 5' AGGAGATATACCATGGATGGCCAAAAAAACCAGCAGCA 3'
	Reverse - 5' GGTGGTGGTGCTCGAGTGAGTTACGCGGTTTCATCACA 3'
<b>CYP81D11</b>	Forward - 5' CCAGCAGCAAAGGTGAGCTCAGCAGCACCAAAAACC 3'
	Reverse - 5' TGGTGCTACCGGATCCCGGACAACCTTGCATCCAG 3'
<b>CYP81D11-int</b>	Forward - 5' TTTTTGTTGTTAGCAGCCAT 3'
	Reverse - 5' TCAAAACTGAACCAGACTAC 3'
<b>CYP81D11-2</b>	Forward - 5' AGGAGATATACCATGGATGGCCAAAAAAACCAGCAGCA 3'
	Reverse - 5' GGTGGTGGTGCTCGAGCGGACAACCTTGCATCCAGAAC 3'
<b>CYP81A9</b>	Forward - 5' AGGAGATATACCATGGATGGCCAAAAAAACCAGCAGCA 3'
	Reverse - 5' GGTGGTGGTGCTCGAGCAGCGGTTTCAGAACCAC 3'

Each reaction mix consisted of approximately 50 ng of plasmid DNA, 20  $\mu$ M of each primer, 10 mM dNTPs mix, 2 U of *Pfu* DNA polymerase in 10x *Pfu* buffer and water to a final volume of 50  $\mu$ l. The thermocycler was set to perform a first denaturation step at 95 °C, 30 cycles of 95 °C for 30 s (denaturation),  $T_m$ -5 °C for 30 s (annealing), 72 °C for 3 min (extension) followed by a final extension step at 72 °C for 5 min.

The modified genes were then cloned via InFusion (Materials and Methods, Chapter 2.4.6) into the destination expression vector.

The vectors used in this chapter, pET28a and pMA-T are shown in figure 3.4. The *Nco*I and *Xho*I restriction sites flanking the multiple cloning site (MCS) in pET28a, were exploited to clone the S-tagged modified gene of interest, upstream of the C-terminal His-tag.



S-CYP-L sequence	<p>ccatggCCAAAAAACCAGCAGCAAAGGTgagctcCCG AGC ACC GAA  CAG AGC GCA AAA AAA GTT CGT AAA AAA GCC GAA AAC GCC CAT  GGT CCT GGT CCA GGT CCT GGT CAT ATG CGT CTG GCG AGC ACC  CAT ATGg gatccGGTAGCACCAGCAGCGGTAGCGGTcctagg</p>
---------------------	--

**Figure 3.4:** Plasmid maps for pET28a and for the synthetic pMA-T plasmid, carrying the solubility tag (S, highlighted in blue in the sequence reported) and the linkers (highlighted in yellow, purple, grey and green), used for the expression of the target genes. The yellow region corresponds to the *B. megaterium* BM3 natural linker between the heme and the reductase, the purple to a flexible glycine-proline loop, the grey to the linker engineered for the P450cam-RhfRed fusion system. In small case are the restriction sites for the endonucleases: *NcoI*, *SacI*, *BamHI*, *AvrII*. The green region encodes for the ‘lam’ linker (152).

### **3.2.2 Expression trials in *E. coli*: Rosetta 2, BL21(DE3), Arctic express**

For the complete expression procedure, see Materials and Methods, paragraph 2.5.1.1.

After sequencing verification, the constructs were transformed into BL21 (DE3), Rosetta2, A. express *E. coli* strains, using the conventional heat-shock protocol (Materials & Methods, Paragraph 2.4.7.1). Single colonies were picked and cultured in 5-ml batches of LB medium with appropriate antibiotics for selection, shaking at 37 °C overnight. The following morning, the cultures were upscaled to 50-500 ml in Erlenmeyer flasks containing fresh M9 growth medium. At induction time, the cultures were supplemented with aminolevulinic acid, FeCl<sub>3</sub> and, if needed, IPTG. The expression temperatures used differed, ranging from 16 °C to 28 °C. The cells were harvested after 16-48 h using gentle centrifugation and lysed mechanically (sonication or French press) to release the expressed proteins. Soluble and membrane cellular fractions were separated by centrifugation of the cell lysates at 34,000 g for 15 min.

### **3.2.3 Expression trials in yeast (*Saccharomyces cerevisiae* WAT11 modified strain)**

The native sequences of *CYP81D8* and *CYP81D11*, each already cloned into pYEDP60 shuttle vectors (constructs retrieved from the Bruce's group stock), were transformed with the LiAc/SS (lithium acetate salmon sperm-mediated) method by Gietz and Schiestl (169) into the WAT11 *S. cerevisiae* yeast strain. This particular strain was chosen as it was engineered to express ATR1, one of the two natural Arabidopsis P450 reductases (170). Single yeast colonies (from the glycerol stocks) were inoculated in 20 ml of YPGA medium and grown at 30 °C on the orbital shaker (200 rpm). The following day the optical density of the cultures was measured and adjusted to a final OD<sub>600</sub> value of 0.15 with pre-warmed YPGA medium. The cells were incubated again at the same temperature and rotation speed as above, for 3-5 h, in order to allow at least two cell duplications. The cells were harvested by centrifugation at 3,000 g for 5 min, washed with cold sterile water and spun again at

3,000 g for 5 min at 4 °C. The cell pellet, resuspended in 1 ml of cold sterile water, was subdivided into 100 µl aliquots. The fractions were centrifuged briefly and each pellet was incubated with the transformation mix (240 µl of PEG 3500 50% w/v, 36 µl of 1M lithium acetate, 50 µl of pre-boiled salmon sperm DNA, target DNA to a final volume of 360 µl) at 42°C for 30 min, with gentle inversions from time to time. The reaction mixtures were placed on ice for two min and centrifuged at high speed for 30 s. The pellets were washed with 1 ml of sterile water, spun quickly and resuspended again in 1 ml of cold sterile water. A fraction of the transformed cells (1/20 of the volume) was spread onto selective SGI plates, which were then incubated at 30 °C for three to four days.

Following the protocol by Pompon *et al.* (171), single transformants were cultured in rich liquid medium (SGI starter and YPGE main) in baffled flasks at 28 °C with agitation, on an orbital shaker set at 140 rpm. After 24 h of incubation, protein expression was induced by addition of galactose (stock: 200 g/l, volume added: 1/10 of each culture). The cultures were then left shaking for further 16 h at 25 °C.

Finally, the yeast cells were harvested via centrifugation, washed and disrupted mechanically with glass beads. Microsomes were then recovered via centrifugation and homogenised in a Tris-HCl buffer pH 7.5 with ethylenediaminetetraacetic acid (EDTA) and 30% glycerol. To check the effective expression of the target CYPs, microsome solutions were diluted and scanned using the spectrophotometer at wavelengths between 400 and 500 nm, to verify the presence of the 450 nm peak upon reduction with sodium dithionite and carbon-monoxide treatment.

### **Media and buffers for yeast growth and harvest**

YPGA: 10 g/l bactopeptone, 10 g/l yeast extract, 20 g/l glucose, 200 mg/l adenine.

SGI: 1 g/l bactocasamino acids, 7 g/l yeast nitrogen base, 20 g/l glucose, 20 mg/l tryptophane, 20 g/l agar.

YPGE: 10 g/l bactopeptone, 10 g/l yeast extract, 5 g/l glucose. 30 mL ethanol was added after the autoclave sterilisation.

TEK: 50 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8, 100 mM KCl, pH 7.5.

TES: 50 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8, 109 g/l sorbitol, pH 7.5.

TEG: 50 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8, 30% vol/vol glycerol, pH 7.5.

### **3.2.4 Purification**

The purification of the modified CYP constructs was accomplished through Ni-affinity chromatography, exploiting the C-terminal hexa-histidine tag. A phosphate buffer, either sodium or potassium-based, was used throughout the whole process (Materials and Methods, Paragraph 2.5.6), at pH values ranging from 7 to 8, according to the calculated pI for the protein products. The proteins, eluted with gradients of imidazole (0-500 mM) or histidine (0-50 mM), were dialysed overnight against phosphate buffer at the appropriate pH, then concentrated and stored in 20 % v/v glycerol in the freezer at -80 °C.

### **3.2.5 Protein characterisation: UV-Vis spectrophotometry and activity**

Fractions collected throughout the purification process, as well as eluted proteins were subjected to UV-Vis spectrophotometric evaluation (Materials & Methods, Paragraph 2.5.3) and activity assays, in order to assess the correct folding and activity of the product of interest. The reaction mixture for the activity assays consisted of different amounts of purified protein (25 and 750 nM), an equimolar concentration of purified plant P450 reductase, 200 µM of the substrate cinnamic acid, 300 µM of NADPH and an NADPH regeneration system, made of 2 mM glucose 6-phosphate and 3 U of glucose 6-phosphate dehydrogenase. Samples from the reactions were collected at the start of the assay, after 5 h, 10 h and 20 h, with the reaction from each time point quenched with an equivalent amount of cold methanol. After centrifugation, the samples were analysed via HPLC (see complete separation method in Materials and Methods, paragraph 2.5.8).

## **3.3 Results**

### **3.3.1 Sequence analysis and cloning**

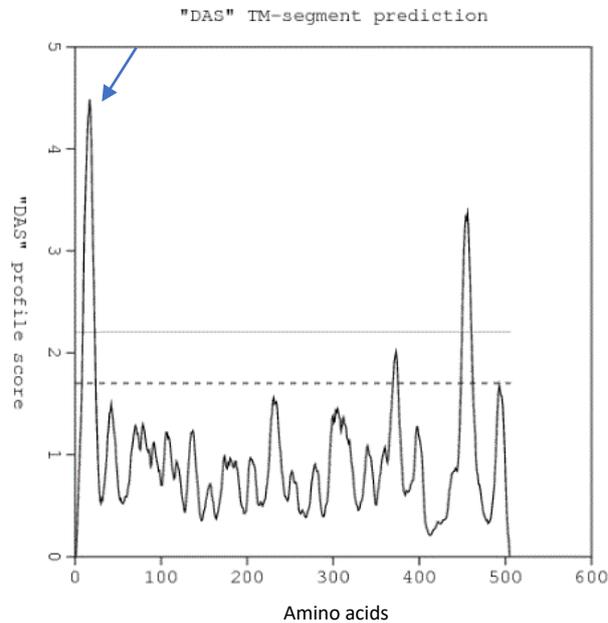
The nucleotide sequences of the target proteins (*AtCYP73A5*, *AtCYP81D11* and *ZmCYP81A9*) were retrieved from the NCBI gene database and translated into the corresponding amino acid sequences with the ExPASy 'Protein Translate' online tool. The obtained sequences were double-checked with the corresponding FASTA files deposited on the UniProt database (172). The verified amino acid sequences were uploaded onto the Das and TMHMM servers for structure prediction (figure 3.5 A-C).

### Chapter 3: Soluble expression of CYPs

**A**

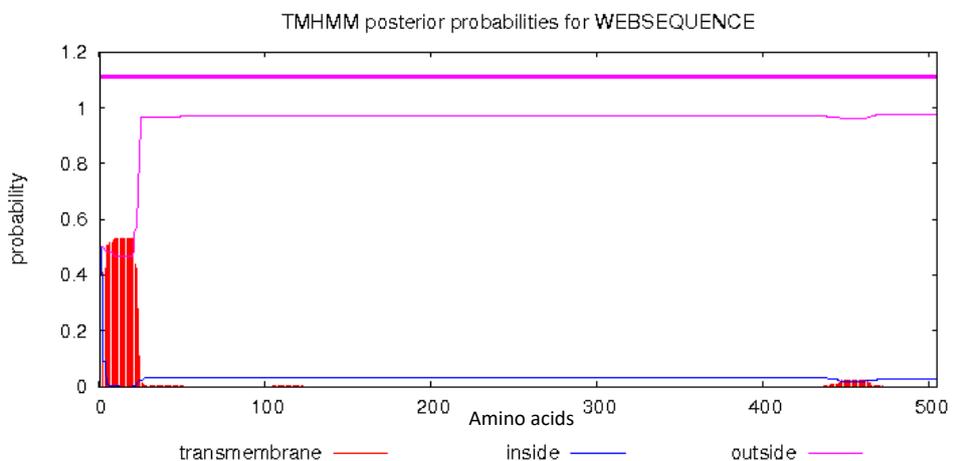
Potential transmembrane segments				
Start	Stop	Length	~	Cutoff
9	24	16	~	1.7
10	23	14	~	2.2
370	375	6	~	1.7
449	462	14	~	1.7
450	460	11	~	2.2

The DAS curve for your query:



**B**

```
# WEBSEQUENCE Length: 505
# WEBSEQUENCE Number of predicted TMHs: 0
# WEBSEQUENCE Exp number of AAs in TMHs: 11.55558
# WEBSEQUENCE Exp number, first 60 AAs: 11.06291
# WEBSEQUENCE Total prob of N-in: 0.50012
# WEBSEQUENCE POSSIBLE N-term signal sequence
WEBSEQUENCE TMHMM1.0 outside 1 505
```



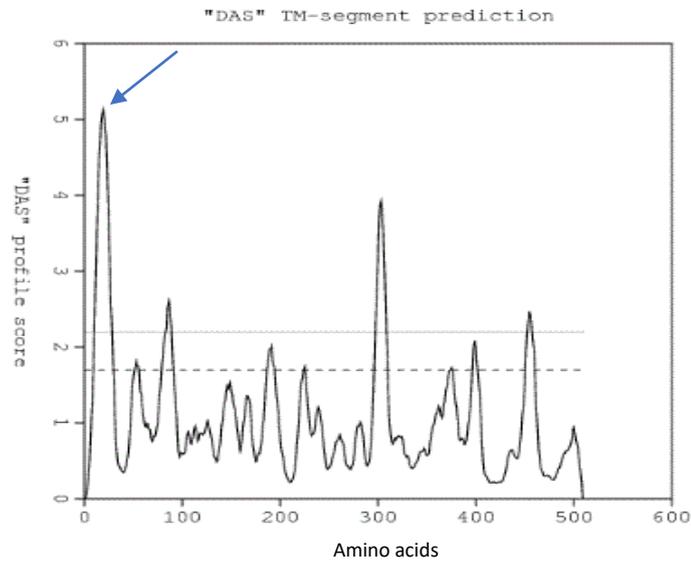
**Figure 3.5A:** Transmembrane domain predictions by DAS (panel A) and TMHMM (panel B) for *AtCYP73A5*. The Das analysis (panel A) predicted that an N-terminal membrane anchor might be formed by amino acids between position 10 and 22. The TMHMM analysis (panel B) did not yield a clear result.

### Chapter 3: Soluble expression of CYPs

**A**

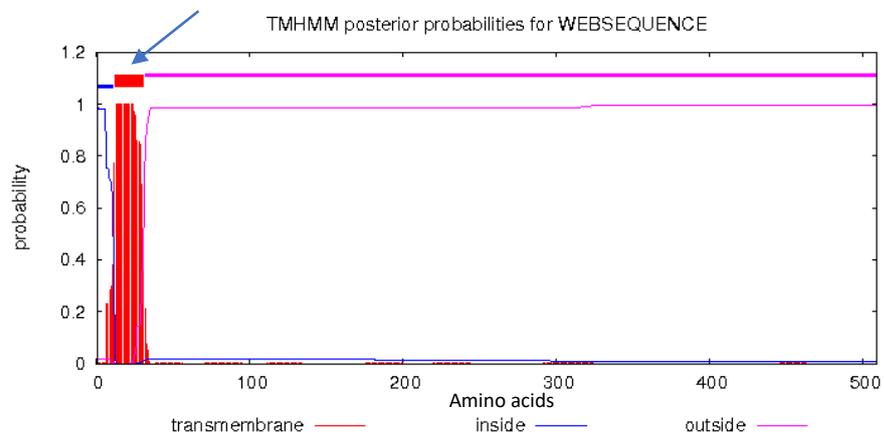
Potential transmembrane segments				
Start	Stop	Length	~	Cutoff
11	28	18	~	2.2
52	55	4	~	1.7
79	98	12	~	1.7
84	89	6	~	2.2
187	194	8	~	1.7
224	225	2	~	1.7
296	309	14	~	1.7
298	308	11	~	2.2
374	375	2	~	1.7
397	401	5	~	1.7
451	460	10	~	1.7
453	457	5	~	2.2

The DAS curve for your query:



**B**

```
# WEBSEQUENCE Length: 509
# WEBSEQUENCE Number of predicted TMHs: 1
# WEBSEQUENCE Exp number of AAs in TMHs: 20.50718
# WEBSEQUENCE Exp number, first 60 AAs: 20.23919
# WEBSEQUENCE Total prob of N-in: 0.98356
# WEBSEQUENCE POSSIBLE N-term signal sequence
WEBSEQUENCE TMHMM2.0 inside 1 11
WEBSEQUENCE TMHMM2.0 TMhelix 12 31
WEBSEQUENCE TMHMM2.0 outside 32 509
```



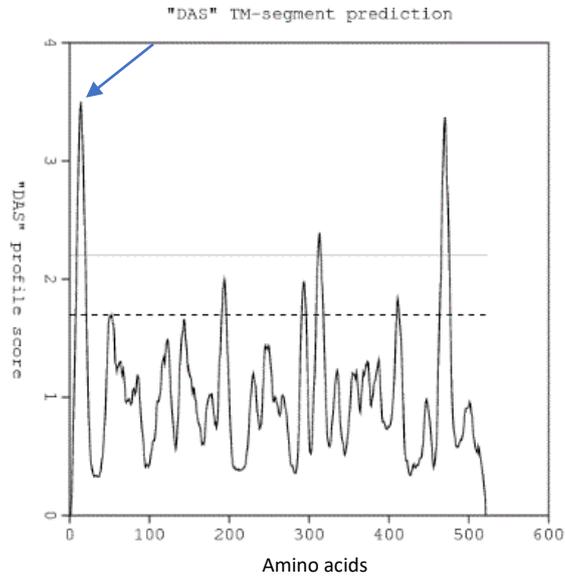
**Figure 3.5B:** Transmembrane domain predictions by DAS (panel A) and TMHMM (panel B) for *AtCYP81D11*. The algorithms identified the region between amino acids 11-28 and 12-31 (as highlighted in blue on each result panel) as the one possibly forming the N-terminal transmembrane domain.

### Chapter 3: Soluble expression of CYPs

**A**

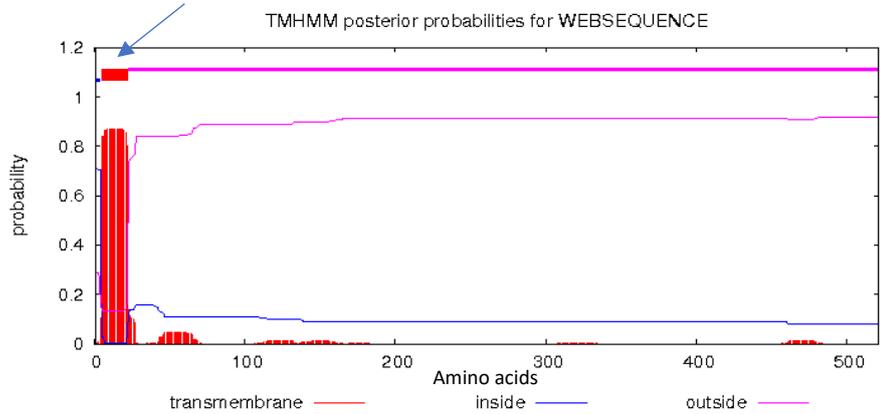
Potential transmembrane segments					
Start	Stop	Length	~	Cutoff	
8	21	14	~	1.7	
9	19	11	~	2.2	
191	196	6	~	1.7	
291	296	6	~	1.7	
309	317	9	~	1.7	
312	314	3	~	2.2	
410	412	3	~	1.7	
464	476	13	~	1.7	
466	475	10	~	2.2	

The DAS curve for your query:



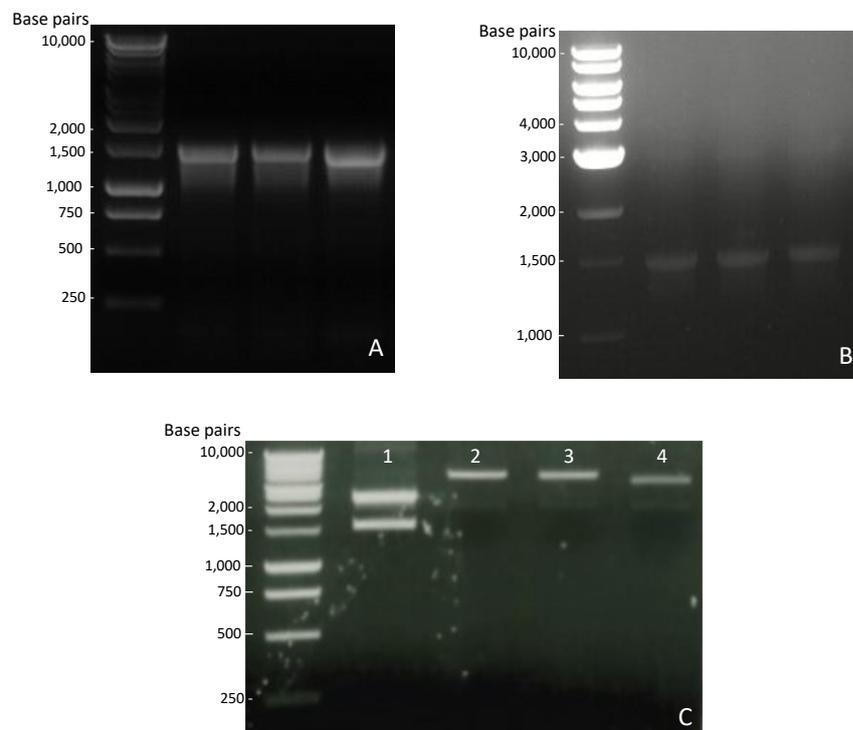
**B**

```
# WEBSEQUENCE Length: 521
# WEBSEQUENCE Number of predicted TMHs: 1
# WEBSEQUENCE Exp number of AAs in TMHs: 17.89786
# WEBSEQUENCE Exp number, first 60 AAs: 16.93067
# WEBSEQUENCE Total prob of N-in: 0.71117
# WEBSEQUENCE POSSIBLE N-term signal sequence
WEBSEQUENCE TMHMM2.0 inside 1 4
WEBSEQUENCE TMHMM2.0 TMhelix 5 22
WEBSEQUENCE TMHMM2.0 outside 23 521
```



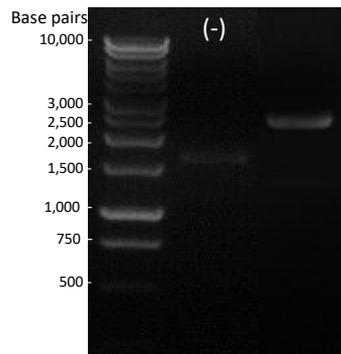
**Figure 3.5C:** Transmembrane domain predictions by DAS (panel A) and TMHMM (panel B) for *ZmCYP81A9*. The two algorithms indicated that the region between amino acids 9-19 and 5-22 might form the N-terminal transmembrane region, as highlighted in blue on each result panel.

As a result of these sequence analyses, the codon-optimised nucleotide sequences of CYP81D11, CYP81A9, were synthesised truncated of the first 30 and 19 codons, respectively (CYP81D11tr, CYP81A9tr). A truncated version of CYP73A5 (CYP73A5tr), lacking the sequence encoding the first 25 codons, was already available in the Bruce group's stock. The nucleotide sequences were further analysed to make sure that there were no internal *SacI*, *Bam*HI, *Nco*I or *Xho*I restriction sites within the sequences of the proteins, as these sites would be used for the following cloning steps. The CYP sequences were amplified using the primers and procedures described in Paragraph 3.2.1. The first cloning step was the PCR-amplification of CYP73A5tr and CYP81D11tr, with primers specifically designed with the flanking *SacI* and *Bam*HI restriction sites, in order to introduce these products into the pMA-T construct, enabling the addition of the S-tag at the N-terminus of the protein (figure 3.6). The sequence for CYP81A9tr was directly excised from the synthetic vector pMK (from the manufacturer) using the *SacI* and *Bam*HI flanking restriction sites.



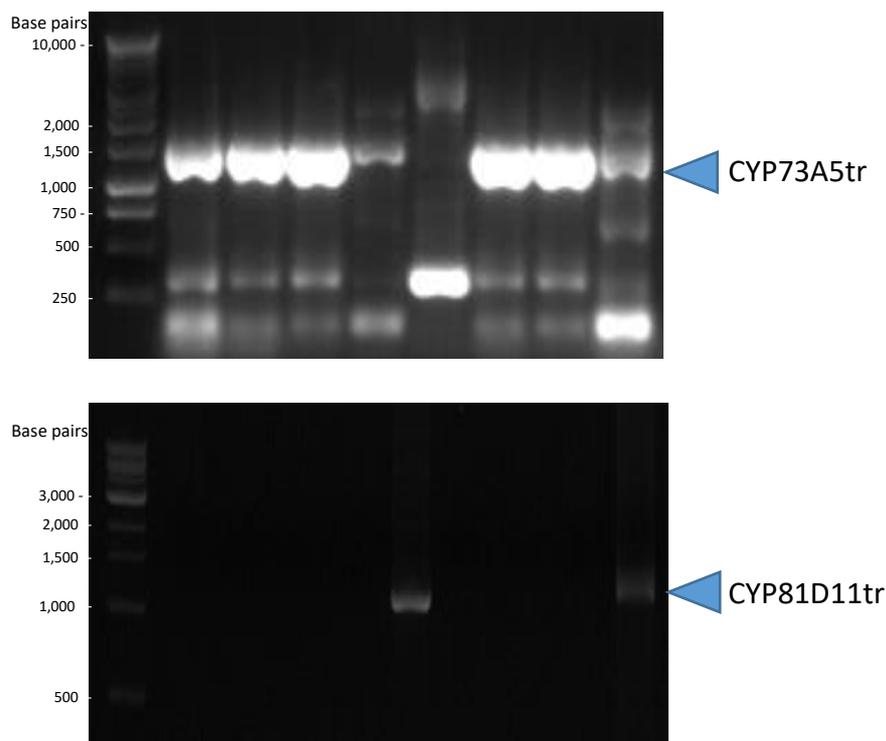
**Figure 3.6:** Electrophoretic analysis on 1.2 % agarose gel of the PCR amplification of CYP73A5tr (A) and CYP81D11tr (B), both performed in triplicate, with *Pfu* polymerase in order to introduce the InFusion-compatible ends for the cloning in pMA-T. (C): Excision of CYP81A9tr from the original synthetic vector (1: double-digested plasmid, 2: *SacI* single cut, 3: *Bam*HI single cut, 4: uncut plasmid). All three products of interest (amplicons for A and B and insert in C) were expected to have a final size of ~1,500 bases.

In parallel, the pMA-T vector was double-digested with the *SacI* and *Bam*HI restriction endonucleases (figure 3.7).



**Figure 3.7:** Double digestion of the pMA-T vector with *SacI* and *Bam*HI. (-) identifies the lane where the undigested plasmid (as negative control) was loaded.

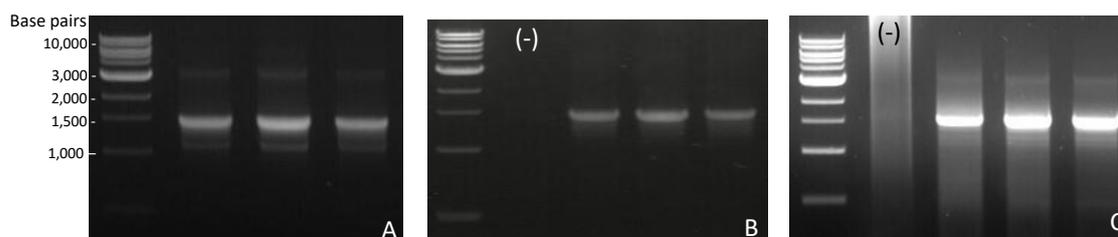
The PCR products were each cloned into the linearised vector via InFusion cloning (S-CYP73A5tr and S-CYP81D11tr) and ligation (S-CYP81A9tr) (for the complete procedure see Materials and Methods, Paragraph 2.4.6). Subsequently, colonies were screened via colony PCR (Materials and Methods, Paragraph 2.4.1), to verify the presence of the desired CYP insert, employing as primers the same oligonucleotides used for the first PCR amplification for CYP73A5tr or internal primers for CYP81D11tr (figure 3.8).



**Figure 3.8:** Colony PCR verification of the presence of the insert in the transformed colonies. The PCR was performed with *Pfu* DNA polymerase using single colonies as templates and the CYP73A5-1 primers (top panel) or the CYP81D11-Int primers (bottom panel).

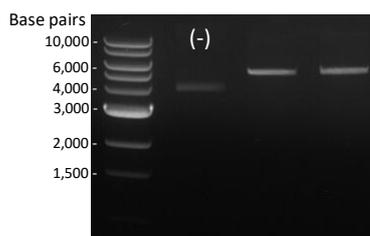
The positive colonies were cultured for subsequent isolation of the plasmid (see Materials and Methods, Paragraph 2.4.5) and sequencing.

The pMA-T\_CYP plasmids which had the correct sequence, comprising the upstream solubility tag, were again subjected to PCR amplification, (figure 3.9) this time employing primers carrying the adapters for the InFusion cloning of the construct into the final expression vector (primers in table 3.2, set 2).



**Figure 3.9:** Electrophoretic analysis of the second set of PCR amplifications of S-CYP73A5tr (A), S-CYP81D11tr (B) and S-CYP81A9tr (C), all performed in triplicate. The same DNA ladder was used for the three gels (NEB 1kb DNA ladder #N3232).

Concomitantly, pET28a was linearised with *Nco*I and *Xho*I (figure 3.10)



**Figure 3.10:** Double digestion of pET28a with *Nco*I and *Xho*I. (-) identifies the lane where the undigested plasmid, as the negative control, was loaded.

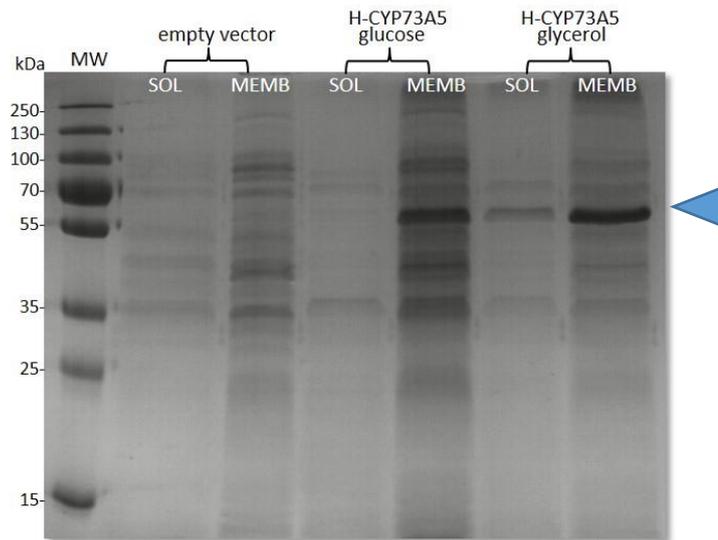
The CYP amplicons were cloned into pET28a and the InFusion reactions were then transformed into competent DH5 $\alpha$  cells. The positive transformants were then inoculated for subsequent plasmid isolation and sequencing.

### 3.3.2 Expression trials

Preliminary expression trials were conducted on S-CYP73A5tr, to establish an expression method for all three CYPs. The presence of the S-tag on CYP73A5tr was predicted to increase the solubility of the CYP. To evaluate the effect of the S-tag, the N-terminally modified S-CYP73A5tr cloned into pET28a was expressed in a range of different bacterial hosts. The outcome of the expression trials can be observed in the SDS-PAGE gel analyses reported in figures 3.11-3.12.

In the first expression screen, BL21(DE3) was chosen as host of expression. The transformed cells were grown in LB and subsequently in M9 medium, induced with 1 mM of IPTG, incubated shaking at 16 °C and harvested after 20 h of incubation. Additionally, a colleague (Dr. Claudia Spandolf, personal communication) who worked on fungal P450s, observed higher yields of expression of P450 constructs when glycerol, instead of glucose, was used as carbon source for the cultures. For this reason, parallel cultures in M9 with 10 mM glucose and in M9 with glycerol (0.4 % v/v) were also carried out in the BL21(DE3) trials, in order to compare the yields.

### Chapter 3: Soluble expression of CYPs

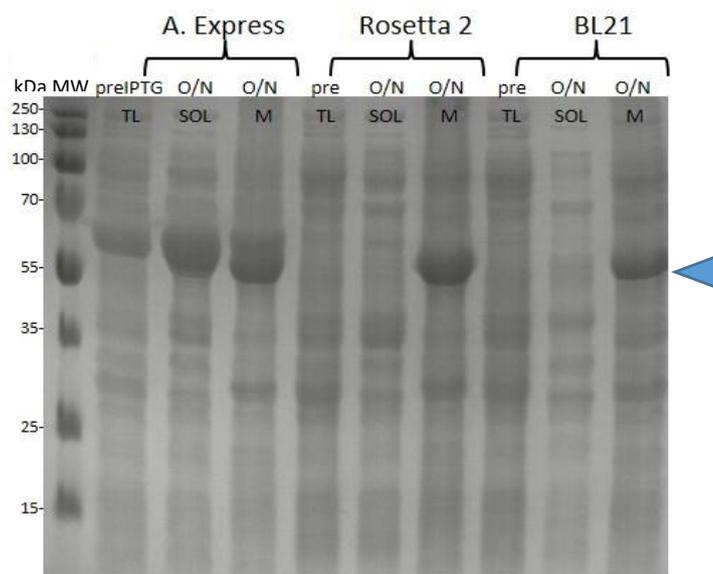


**Figure 3.11:** SDS-PAGE analysis of the expression of S-CYP73A5tr in glucose or glycerol-fed BL21 cultures, compared to cultures harboring the empty vector, in the soluble and membrane fractions of the sonicated cells. First 2 lanes: empty vector / 3<sup>rd</sup>-4<sup>th</sup> lane: glucose-fed batch / 5<sup>th</sup>-6<sup>th</sup> lanes: glycerol-fed batch. The blue arrow pointer indicates the product bands at the correct expected size of 57 kDa.

Bands at the correct size (~ 57 kDa) could be visualised on the gel (in the lanes related to both glycerol and glucose cultures) but no cross-reactivity with the antibody anti-His in western blot assays could be detected. A 57 kDa band was sent for MALDI-TOF MS-MS analysis and identified as the Arabidopsis cinnamate hydroxylase. However, since the C-terminal His-tag could not be detected on the western blot, the protein could not be purified via His/Ni-affinity chromatography.

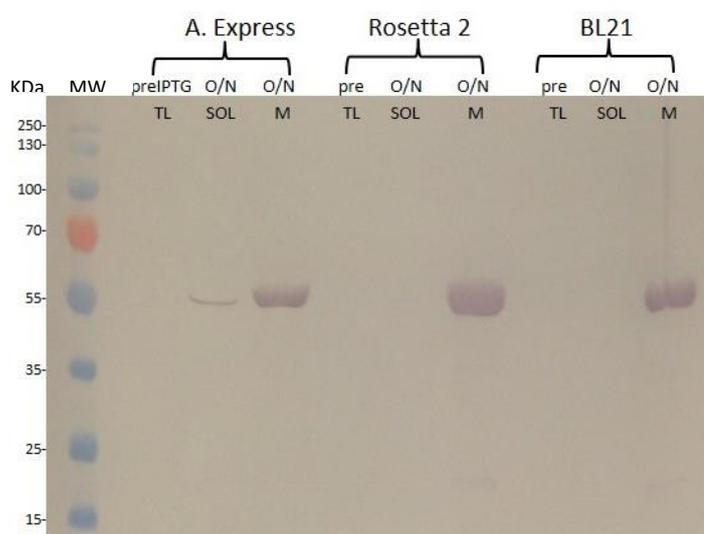
The S-CYP73A5tr construct was subsequently transformed in other *E. coli* strains, in A. express and in Rosetta 2, in parallel with BL21(DE3). Samples before the induction with IPTG (0.5 mM) and at the end of the expression period were collected and analysed via SDS-PAGE (figure 3.12).

### Chapter 3: Soluble expression of CYPs



**Figure 3.12:** SDS-PAGE analysis of the expression of S-CYP73A5tr in three different expression hosts, *E. coli* A. express, Rosetta 2, BL21. For each culture three samples were collected and loaded on the gel: the total lysate pre-induction (pre or preIPTG TL) and the soluble (SOL) and membrane (M) fractions at the end of the expression period, after the cell lysis. The blue arrow pointer indicates the product bands at the correct expected size of 57 kDa.

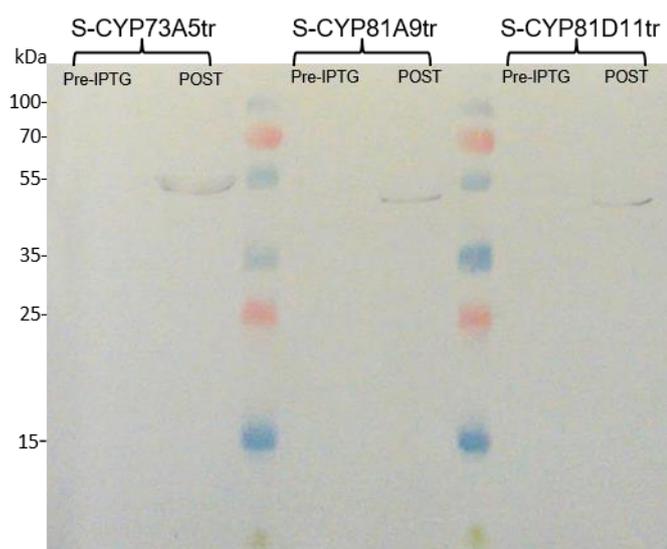
A western blot with antibodies directed towards the C-terminal His tag present in S-CYP73A5tr was performed, in order to verify if the bands with a size between the 55 and 70 kDa markers were effectively given by the protein constructs of interest (figure 3.13).



**Figure 3.13:** Western blot anti-His analysis of the samples collected in the expression trials of S-CYP73A5tr in the three *E. coli* strains (as for figure 3.12). preIPTG TL = total lysate prior to induction; SOL = soluble fraction; M = membrane fractions; pre = before induction; O/N = at the end of the expression period. Expected molecular weight of the products: ~57 kDa.

The western blot analysis (figure 3.13) clearly showed that all the chosen *E. coli* hosts were able to express the 57 kDa His-tagged product in the membrane fraction, but the Arctic express host also showed that the CYP product was present in the soluble cell fraction.

Following the preliminary expression trials with S-CYP73A5tr, all the S-CYPtr constructs (S-CYP81A9tr and S-CYP81D11tr) were transformed into Arctic express competent cells for expression. The outcome of the expression trial was verified via SDS-PAGE and western blot (figure 3.14).



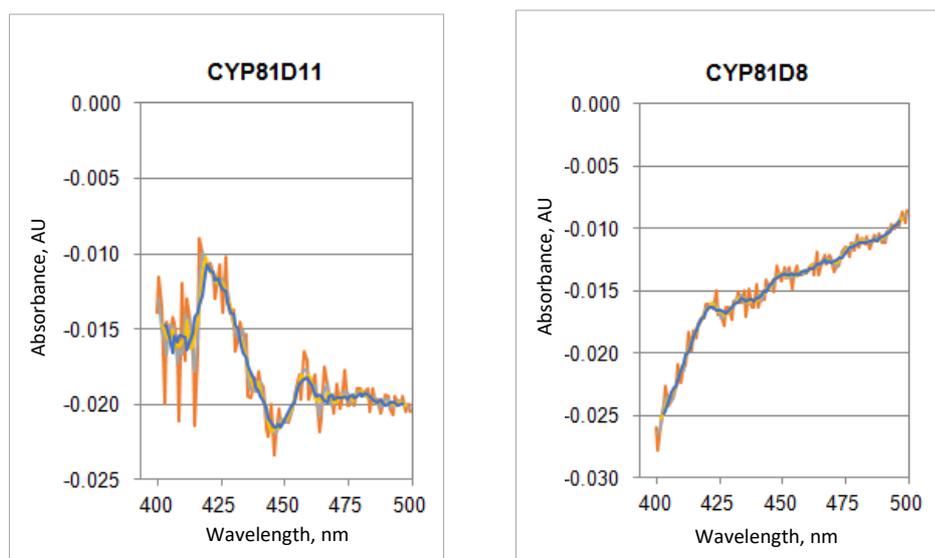
**Figure 3.14:** Western blot analysis of the expression trials of the three S-CYPtr constructs in *E. coli* Arctic express. For each culture a sample of the soluble fraction prior to induction (pre-IPTG, negative control) and another sample of the soluble fraction at the end of the expression period (POST) were collected and loaded on the SDS-PAGE gel, which was then blotted on nitrocellulose and incubated with the antibody for detection. Expected molecular weight of the products: ~55-57 kDa.

The western blot analysis clearly showed that all the S-CYPtr constructs were successfully expressed in the soluble fraction of the Arctic express cells.

### 3.3.3 Expression in yeast

Aiming at acquiring some experience in protein expression using a yeast system, two constructs, pYEDP60-CYP81D8 and pYEDP60-CYP81D11, were transformed into competent *S. cerevisiae* WAT11 yeast cells. Cell microsomes were then harvested and analysed using the UV-Vis spectrophotometer (see Materials and Methods,

Paragraph 2.5.3). Figure 3.15 shows the CO-binding spectra of the extracted microsomes, used to verify the presence of P450 products.

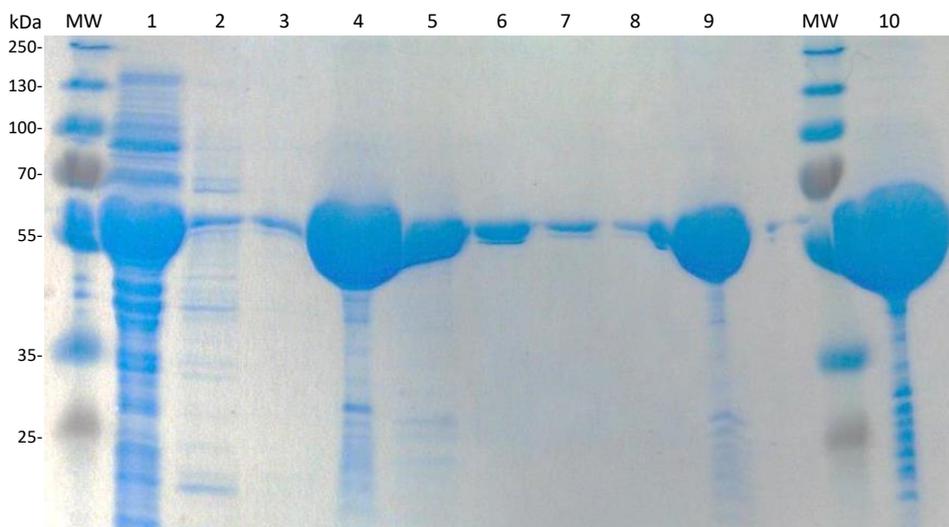


**Figure 3.15.** CO-binding spectra of the microsomes collected from yeasts transformed with pYEDP60-CYP81D11 (left panel) and pYEDP60-CYP81D8 (right panel). Microsomes were obtained via mechanical lysis of the yeast cells, separated from the residual debris via multiple centrifugations and homogenised in Tris-HCl buffer pH 7.5.

No evidence of CYP expression could be observed, as the signature 420/450 nm peaks were both missing from the microsomes' CO-binding spectra.

### 3.3.4 Purification of the S-CYP73A5tr construct

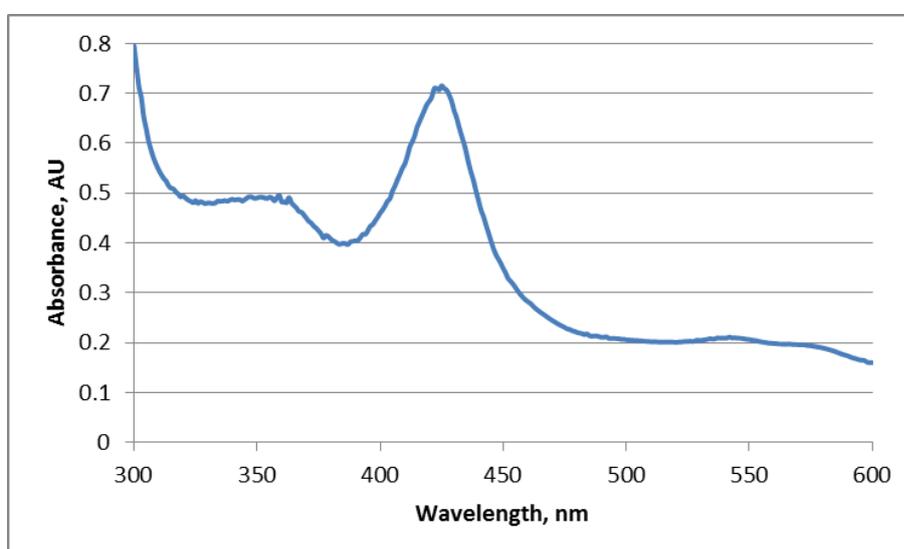
The soluble fractions containing the S-CYP73A5tr product from Arctic express cells were filtered and subjected to Ni-affinity chromatography (See complete procedure in Materials and Methods, Paragraph 2.5.6). The elution, carried out using a gradient of imidazole, from 10 to 300 mM, is shown in figure 3.16.



**Figure 3.16:** SDS-PAGE analysis of the samples collected during the purification of S-CYP73A5tr. Lane 1: soluble fraction pre-loading, 2: wash, 3-8: elutions, 9: post-dialysis, 10: concentrated, MW: molecular weight marker. Expected molecular weight of the product: 57 kDa.

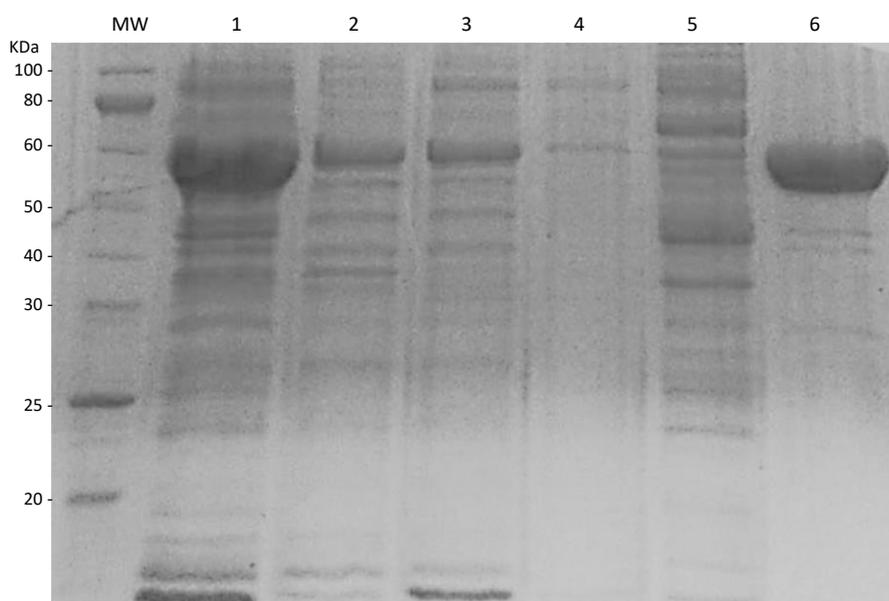
The S-CYP73A5tr protein was successfully purified with a yield, calculated via Bradford assay (see Materials and Methods, paragraph 2.5.7), of 5 mg, from a 200 ml starting culture.

When examined using the UV-Vis spectrophotometer, the product showed the characteristic absorbance peak at 420 nm related to the P450 enzyme in the resting state (figure 3.17).



**Figure 3.17:** UV-Vis absorbance spectrum of the final purified S-CYP73A5tr.

An alternative purification procedure was trialed, employing a gradient of histidine, from 5 to 50 mM for the final elution, in place of imidazole, which has been reported as an inhibitor of CYP enzymes. As shown in the resulting SDS-PAGE gel (figure 3.18), the S-CYP73A5tr can be isolated from the cell extract in the absence of the commonly used imidazole (which could affect the activity of the protein), simply exploiting an excess of histidine to displace the protein-column interactions.



**Figure 3.18:** SDS-PAGE analysis of the samples collected during the purification of S-CYP73A5tr. Expected molecular weight of the product: 57 kDa (Lane 1: pre-loading, 2: unbound, 3-5: wash, 6: final elution, MW: molecular weight marker).

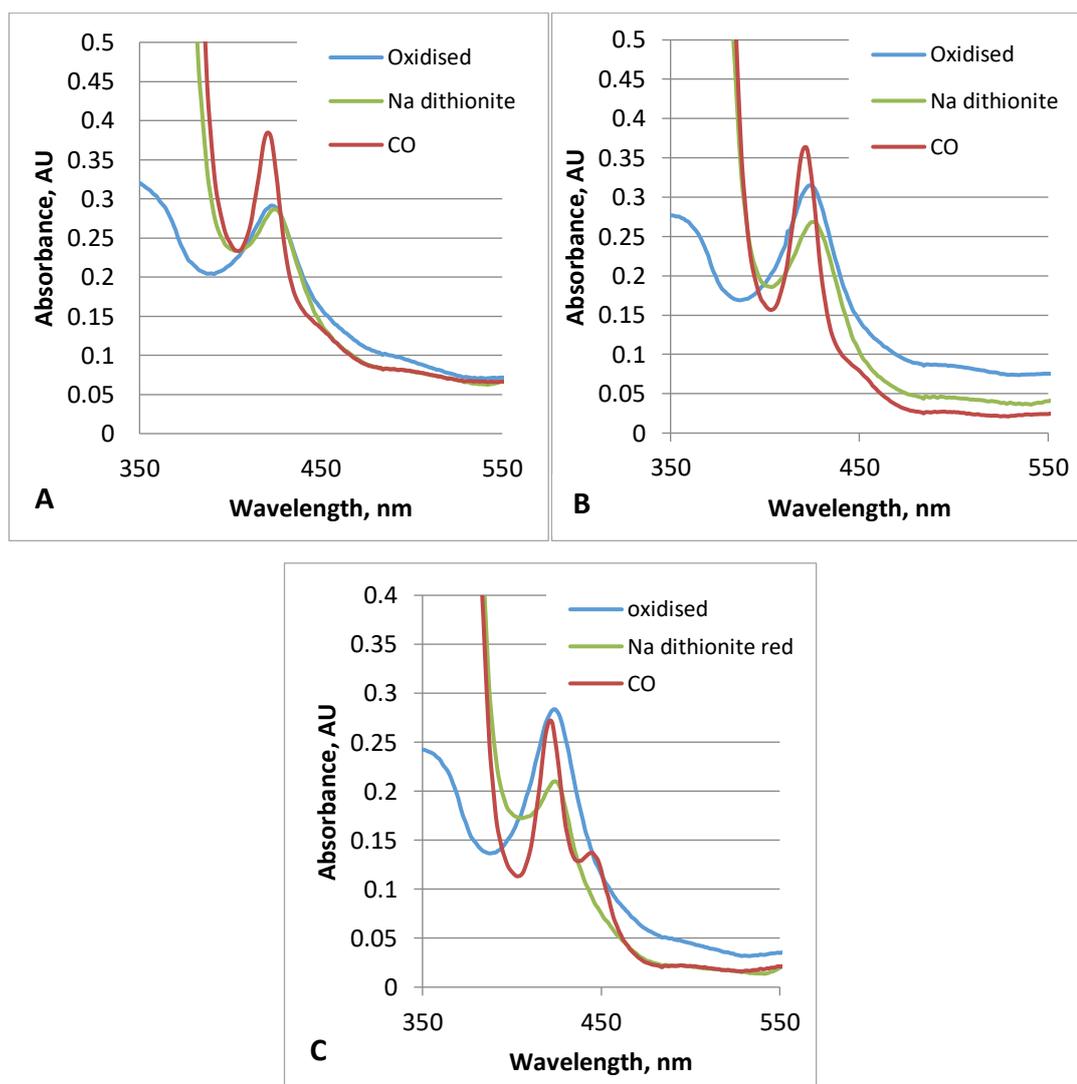
Interestingly, the protein elution fractions showed a violet colouration (figure 3.19).



**Figure 3.19.** Picture of the cuvettes containing only phosphate buffer (blank, left) and the eluted S-CYP73A5tr (right).

The same purification procedure was employed for S-CYP81D11tr, but in this case, even modifying the composition of the lysis, binding and elution buffers, the purification was unsuccessful.

Both purified S-CYP73A5tr batches were subjected to UV-Vis spectrophotometric characterisation, to assess the correct folding of the product. A purified sample of the microbial P450 XplA, kindly donated by Dr. Dana Sabir, was used as a positive control (figure 3.20).



**Figure 3.20:** CO-binding spectrophotometric characterisation of the purified S-CYP73A5. A: batch purified with imidazole, B: batch purified with histidine, C: XplA control.

In the CO-binding assay there was no shift in the absorbance peak from 420 nm (related to the protein in the resting state) to 450 nm upon binding of the reduced product with carbon monoxide. The XplA positive control consisted of two protein populations: part showed a peak at 450 nm, but most retained the 420 nm peak. These results suggest that either most of the XplA control is made of non-correctly folded proteins or the CO-bubbling step was not sufficient to induce the desired structure reconfiguration homogeneously in the whole sample.

### 3.3.5 Activity assay

The CO-binding spectra are not always the only and unconfutable proof of active and correctly folded CYPs, as it has been demonstrated that some active plant CYP enzymes are able to catalyse reactions in absence of a “correct” CO-binding shifted spectrum (personal communication). For this reason, activity assays were performed with the purified S-CYP73A5tr towards the natural substrate, cinnamic acid, supplemented with a purified P450 reductase (ATR2 from Arabidopsis, cloned and expressed separately using the same system as for the CYP enzyme), NADPH as the electron donor and an NADPH regeneration system.

Upon HPLC analysis of the samples collected from the two sets of reactions, with 25 nM and 750 nM of S-CYP73A5tr, there was no significant decrease of the amount of substrate, cinnamic acid, nor formation of product, *p*-coumaric acid, within the chosen time frame (20 h) (Data not shown).

## 3.4 Discussion

The biochemical and structural characterisation of eukaryotic P450 enzymes has been significantly constrained by the challenge of producing these enzymes in conventionally exploited bacterial systems. The situation is particularly problematic when plant CYPs are taken into consideration: almost 14,000 P450 genes have been annotated so far (according to the latest count, carried out on 6 April 2016 (173)), but most of the encoded proteins still remain uncharacterised. In this chapter, several

expression trials with three different plant P450 targets were carried out, with the goal of designing an optimal expression protocol to obtain soluble and active P450 products.

### **Gene expression in bacteria**

The process started with the optimisation of the original nucleotide sequence, in order to ease the translational process in the bacterial host of expression. The N-terminus membrane-binding domains, determined via online sequence analysis tools, were cleaved and replaced with a short peptide, S- (MAKKTSSKG), which, as reported by many researchers in the cytochrome P450 field, helped in the expression of other human and plant P450 enzymes. The engineered constructs of S-CYP73A5tr, S-CYP81D11tr and S-CYP81A9tr were each cloned into His-tagged pET28a expression vectors and were transformed into multiple *E. coli* hosts. A successful soluble protein expression was obtained for all three protein targets when transformed *E. coli* Arctic express cultures were grown shaking in M9 medium at low temperatures (16 °C). S-CYP73A5tr was also successfully purified via Ni-affinity chromatography, using either a gradient of imidazole or histidine. When examined using the spectrophotometer, the purified proteins presented the signature absorbance peak at 420 nm (typical of P450 enzymes in the oxidised resting state), but they did not show the shift of the absorbance peak to 450 nm, which should be expected for correctly folded P450 proteins, when reduced and bound to carbon monoxide. This feature might be explained by the fact that the removal of the N-terminal domain and/or the addition of the 'S' peptide might have affected the overall folding of the protein product.

In assays measuring the activity of S-CYP73A5tr, there was no significant hydroxylation of the provided substrate, cinnamic acid, to *p*-coumaric acid. This result might depend on the chosen set up: a different medium, different temperature or higher concentrations of the catalyst or different ratios of P450:P450 reductase might be required to carry out the reaction. The activity may be enhanced by the addition, in the reaction mix, of cyt<sub>b5</sub>, as reported for assays with purified human CYP2A6 (174), human CYP17A1 (175), fungal CYP5136A1 (176). As Schenkman and Jansson

explained,  $\text{cyt}_{b5}$  plays a direct role in the electron flux, by transporting one of the electrons or facilitating the interaction between the reductase and the CYP (177).

Possible solutions, which have worked in the past for other eukaryotic P450 enzymes, is the co-expression of the redox partner or the creation of fusion polypeptides. In the latter case, the CYP and the reductase are encoded as a single macroprotein, mimicking the naturally-occurring and highly active P450 BM-3 (CYP102A1) from *B. megaterium*.

### **Gene expression in yeast**

Yeasts constitute valuable alternatives to *E. coli* hosts for recombinant protein expression, given the relatively low costs for the cultures, and the presence of eukaryotic post-translational modification systems, which are particularly useful when the protein targets are of eukaryotic origin. WAT11 is a modified *S. cerevisiae* yeast strain, engineered to co-express the ATR1 P450 reductase from Arabidopsis. Several plant CYPs, cloned in the shuttle vector pYEDP60 (which contains a galactose-inducible GAL10-CYC1 promoter) have been expressed using the yeast WAT11 and WAT21 strains, created by Pompon *et al* (171). Among these CYPs are wheat (*Triticum aestivum*) CYP51 (178), soybean (*Glycine max*) CYP93A1 (179), Jerusalem artichoke (*H. tuberosus*) CYP76B1 (180) and CYP81B1 (181) and Arabidopsis CYP88A3 and CYP88A4 (182). The expression trials of the native sequences of *At*CYP81D11 and *At*CYP81D8 (cloned in the pYEDP60 vector) were not successful, as the spectrophotometric analysis of the extracted cell microsomes did not display the expected 420 nm and 450 nm signature absorbance peaks for CYP enzymes. This result might derive from the different codon usage between the source of the CYP sequences (Arabidopsis) and the host of expression, as the sequences were not optimised prior to cloning into the shuttle vector. Additional factors affecting the production of the target proteins could be the instability of the recombinant products or the presence of endogenous proteolytic systems.

In the context of heterologous protein expression of P450 enzymes in yeast, *Pichia pastoris* (183-185) might be a valuable alternative species to *Saccharomyces*. Several

groups have already employed *Pichia* for the production of functional plant CYPs: among these are CYP79D1 and CYP79D2 from Cassava (*M. esculenta*) (186) and Arabidopsis CYP85A2 (187).

# Chapter 4: Expression screening and characterisation of P450-reductase fusion proteins

## 4.1 Introduction: plant P450-reductase fusions

CYP-mediated biotransformations are constrained by several drawbacks, such as the requirement of an electron donor, the necessity of flavin-based cofactors or iron-sulfur clusters, and the need for a redox partner for an efficient electron shuttling and catalysis. Another aspect that has to be taken into account when working with CYPs is the high occurrence of undesired uncoupled reactions, which lead to the formation of reactive oxygen species and heavily affect the catalytic efficiency of these enzymes (43, 188).

To overcome the problem of the low efficiency of electron transport from the electron donor to the P450 heme, fusion systems, mimicking the widely studied *B. megaterium* P450-BM3 (CYP102A1) (figure 4.1), have been developed. In these artificial fusions systems, the CYP is directly connected to the redox partner, sometimes using a synthetic linker as a spacer, creating a single polypeptidic chain in place of two different proteins.



**Figure 4.1:** Structural configuration of *B. megaterium* P450-BM3.

The resulting fusion system displays a more efficient electron transport, which thus results into improved enzyme activity. Murakami *et al*, in 1987, were among the pioneers in the production of P450-reductase fused proteins, generating a 130 kDa fusion of a rat CYP (P450c) with a rat reductase in yeasts. This particular product exhibited an activity four-fold higher than that of the native non-fused P450 (189).

Alternatively, an increased electron transport could be achieved using reconstituted systems, whereby purified P450 reductases are incubated with the P450 enzyme of interest. An example is the use of purified rat liver microsomal NADPH P450 reductase incubated with bovine P45017 $\alpha$ , heterologously expressed in *E. coli* cells (136).

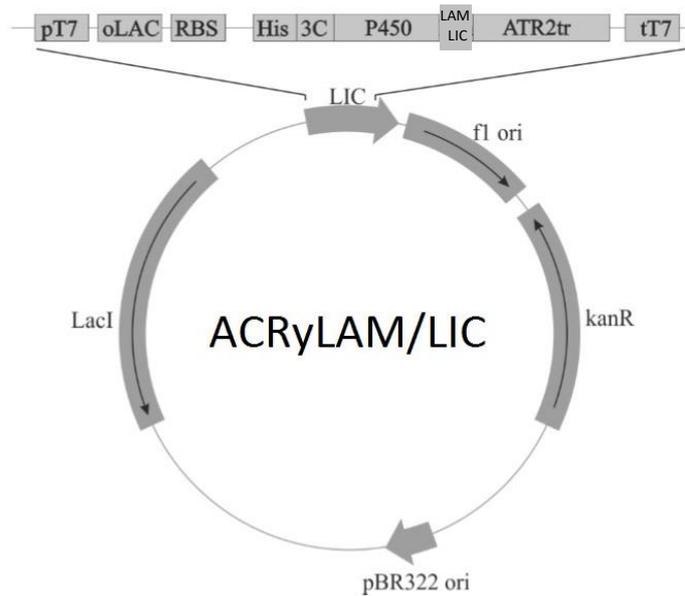
One of the first artificial plant fused systems, developed in 1995 by Hotze and coworkers, was the fusion of a Madagascar periwinkle (*Catharanthus roseus*) cinnamate hydroxylase (C4H) to a P450 reductase, from the same organism. The fusion system was built using traditional digestion/ligation cloning: the C4H sequence was truncated of its aminoterminal membrane-anchoring domain (at different lengths up to its total deletion, for comparison purposes) and introduced in the vector containing the N-terminally-truncated reductase, leading to six different final translational fusions. The linker was a short string of amino acids, with alternating serines and glycines (sequence: STSSGSG). Following expression in *E. coli*, the activity assays demonstrated that almost all the developed plant fusion chimeras were active towards the given substrate (190). Later, the same group generated another plant translational fusion, comprising CYP71D12 and a reductase, both from *C. roseus*, which allowed the verification of the involvement of this CYP in the biosynthesis of bisindole alkaloids (191). In 2002 Didierjean *et al.* created two chimeras of Jerusalem artichoke (*H. tuberosus*) CYP76B1 with the reductase from the same organism, HTR1; 76B1-HTR1 contained the full-length reductase, whereas 76B1-HTR1 $\Delta$ 57 included an N-terminal truncated reductase. The aim of this study was to create an efficient system for the metabolism of phenylurea herbicides. In this case, it was demonstrated that the higher efficiency could be rather achieved with the CYP alone, not fused to the reductase (192). Leonard *et al.* developed a fusion between a *G. max* isoflavone synthase (CYP93C) and a *C. roseus* P450 reductase, connected via a polyGlySer linker (the same as for the *C. roseus* fusions mentioned above, with an additional upstream glycine). The CYP93C was N-terminally modified by trimming, or replacing, the membrane-spanning domain, generating different constructs for activity and solubility comparisons. The highest activity towards the given substrates, naringenin or liquiritigenin, was achieved when the fusion contained the truncated CYP93C, with an N-terminal 'MALLLAVF' peptide, from bovine P45017 $\alpha$  (see Chapter 3,

Introduction section) (193). An additional example of an effective amino terminal modification is the introduction of the hydrophilic sequence MAKKTSSKG, originating from the human CYP2C3 (141), which allowed the soluble expression of CYP73A5tr, CYP81D11tr and CYP81A9tr, as discussed in Chapter 3.

Another aspect of particular importance in the development of such fusion systems is the linker between the P450 and the reductase partner. This spacer should be appropriately designed in order to gain the optimal distance and orientation of the two units for catalysis to occur (194). In bacteria, Robin *et al.* generated several fusions of the camphor-hydroxylating P450cam from *P. putida* with the *Rhodococcus* Rhf reductase (RhfRed), varying the length (2-16 amino acids) of the natural linker in the RhF fused system (195).

Dr. Sabbadin, from Prof. Neil Bruce's research group, exploited the same RhfRed domain to create a P450/reductase fusion platform, called LICRED. This platform was validated with the explosive-degrading XplA (CYP177A1) from *Rhodococcus* and P450cam (CYP101A) from *P. putida*: in both cases the LICRED construct led to the production of soluble and active P450 fusions. The LICRED fusion construct enabled the functional characterisation of additional bacterial CYPs: a set of *N. farcinica* P450 enzymes (90, 91), as well as the identification of pharmacologically-interesting P450s from *Rhodococcus jostii* (196).

The ACryLIC and ACryLAM constructs are two artificial chimeric systems developed previously, in a collaboration between Prof. Neil Bruce's and Prof. Gideon Grogan's research groups, by Dr. Schuckel: target P450s were fused to the Arabidopsis ATR2 P450 reductase with two different linkers (RAFSS for the LIC system and GSTSSGSG for the LAM platform) and cloned into a pET28 vector (figure 4.2).



**Figure 4.2:** Map of the ACryLAM/LIC fusion platform, designed by Dr. Schuckel. The his-tagged N-terminally truncated P450 is connected through a LAM (sequence: GSTSSGSG) or LIC (sequence: RAFFSS) linker to the N-terminally truncated ATR2 reductase from Arabidopsis.

The starting point of Dr. Schuckel's project was the identification of an appropriate host for the expression of chosen plant P450 targets (Arabidopsis CYP81D11, CYP81D8 and CYP89A9, as well as the peppermint CYP71D15 PM2-2), by comparing different *E. coli* strains with WAT11/WAT21 modified *S. cerevisiae* (171). When yeasts were employed for expression, no CO-shift (420 - 450 nm) could be observed in the yeast microsomal preparations, indicating that none of the plant P450 targets was correctly folded. On the other hand, CYP81D11 and CYP81D8 were expressed in *E. coli* Rosetta 2, grown on LB medium, but despite the presence on the SDS-PAGE gel, the products could not be detected via western blot anti-histidine tags. Subsequent work focused on the choice of the reductase partner. Urban *et al.*, in a function-complementation study carried out in *S. cerevisiae*, identified two cDNAs from an Arabidopsis expression library, encoding for proteins that were able to compensate the disruption of an endogenous P450 reductase in the yeast host. It was hypothesised that these products were involved in the same type of reactions carried out by the original, disrupted, P450 reductase. Sequencing analysis of the isolated cDNA targets, alignments with other known P450 reductases, as well as activity assays with cytochrome c and CYP73A5 demonstrated that these two proteins, could

be effectively classified as P450 reductases and therefore were named *Arabidopsis thaliana* Cytochrome P450 reductase (ATR) 1 and 2 (170). ATR1 (Locus tag: AT4G24520, gene ID: 828554) and ATR2 (Locus tag: AT4G30210, gene ID: Gene ID: 829144) share a 66 % amino acid sequence identity, with conserved FAD/FMN/NADPH binding domains, with the N-terminus being the most diversified part of the amino acidic sequences, due to two different -in sequence and length-signal anchors. In addition, multiple studies demonstrated that if on the one side ATR1 is expressed at constant levels, the expression of ATR2 can be upregulated by wounding, as well as light and low temperature conditions, right before, or in parallel to, the induction of other effectors in the phenylpropanoid pathway (197-199). A third putative reductase has been also identified and named ATR3. Nevertheless, sequence alignments, localisation and functional comparisons to ATR1 and ATR2 showed that ATR3 cannot be classified as a P450 reductase. It was instead proven that ATR3 is a cytosolic diflavin reductase involved in early plant embryogenesis (200).

In Dr. Shuckel's work the ATR1 and ATR2 reductases from *Arabidopsis* were truncated to remove the sequences encoding the amino-terminal membrane targeting domain. ATR1tr and ATR2tr were expressed in *E. coli*, purified and assayed for activity. ATR2tr was chosen due to its higher catalytic activity when compared to ATR1tr. The linker used by Leonard *et al.* (sequence: GSTSSGSG), here named LAM, and the short string of amino acids RAFSS, resulting from the intermediate steps for the ligation independent cloning procedure, named LIC, were then employed to create the artificial fusion constructs between the ATR2tr reductase and two P450 targets: *G. max* isoflavone synthase (IFS) and the *Arabidopsis* CYP73A5. In the CYP73A5-ATR2 fusion, the first 25 amino acids from the CYP were also removed, to decrease the level of hydrophobicity of the final protein. These fusions were then compared in terms of activity and solubility to constructs where the same P450 enzymes were fused to the *Rhodococcus* reductase, RhfRed. The highest activities were observed for the IFS-RhfRed fusion and for the CYP73A5tr-ATR2tr chimera, leading to the conversion of up to 50% of naringenin (IFS substrate) and 30% of cinnamic acid (CYP73A5 substrate). For the plant fusion, CYP73A5tr-ATR2tr, despite the solubility of the ATR2tr reductase

and the truncation of the N-terminal membrane-binding domain of CYP73A5, the activity was only detected in the cell membranes and not in the soluble fractions, indicating that the fusion protein was insoluble (152).

The objective in this chapter is the development of a stable and active fusion construct, with the CYP directly connected, through an appropriate linker, to a redox partner. All the constructs here generated contained the Arabidopsis cinnamate hydroxylase (*AtCYP73A5*), employed as a model P450 also in Chapter 3, and the Arabidopsis ATR2 reductase (*AtATR2tr*).

## **4.2 Materials and Methods used for the expression screening of the plant fusions**

### **4.2.1 Cloning steps**

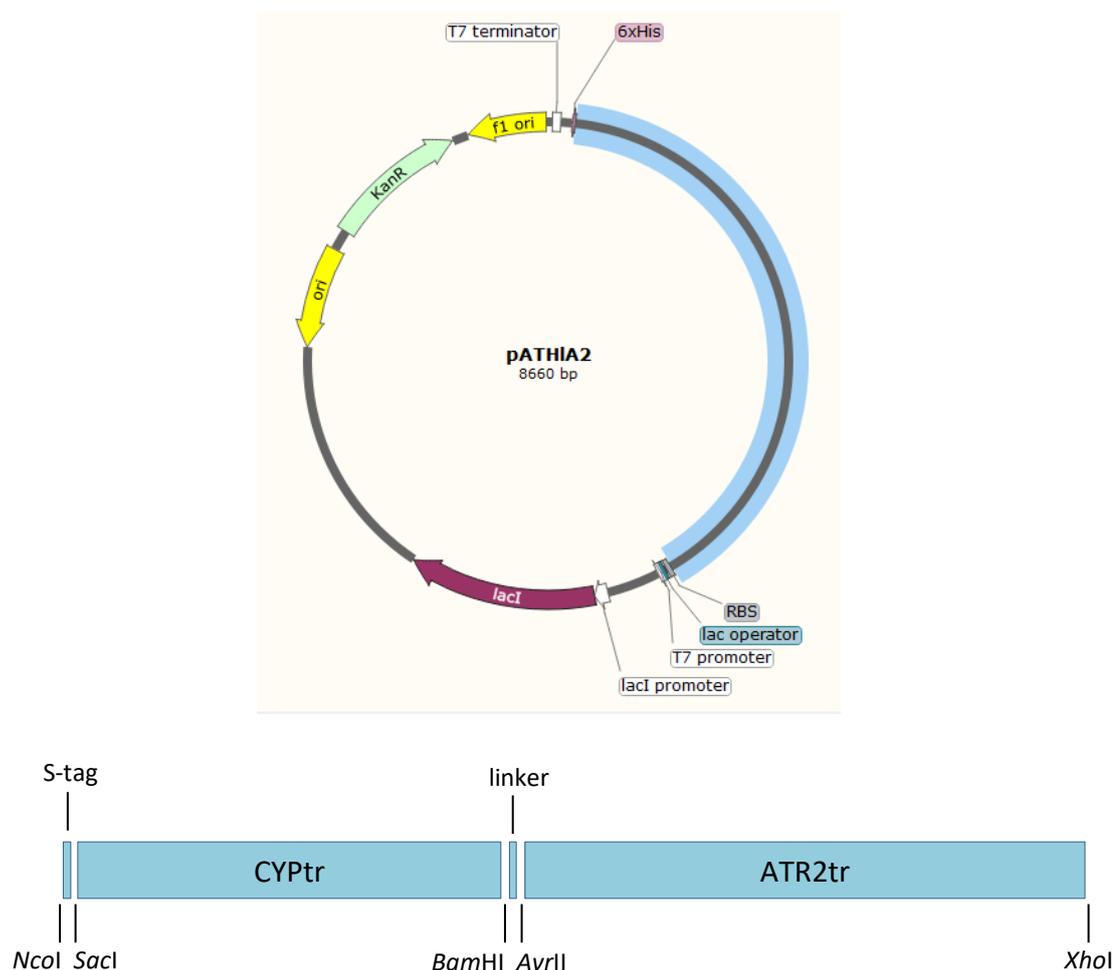
The cloning procedure for the creation of the fusion platform, named pAtHIA2, was similar to the procedure exploited in Chapter 3, for the production of the S-CYPtr constructs. The first step consisted of the introduction of the chosen redox partner, *AtATR2*, in the final expression vector, pET28a, via InFusion cloning (for the complete procedure, see Materials and Methods, section 2.4.6), between the *NcoI* and *XhoI* restriction sites. The sequence of the ATR2 reductase (Gene ID: 829144, Locus tag: AT4G30210) was analysed using the Das (<http://www.sbc.su.se/~miklos/DAS/>) and TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) online transmembrane structure prediction tools, to identify the N-terminal membrane-binding domain of the protein, which had to be cleaved (exactly as for the CYPtr in Chapter 3) to increase the chances of obtaining an overall soluble protein. The oligonucleotides for the PCR amplification of the *ATR2* sequence were specifically designed to allow for the extension and amplification of the N-terminally truncated version of the reductase (*ATR2tr*). The *S-CYP73A5tr* sequence in pMA-T was PCR-amplified using primers specifically designed for the introduction of flanking *NcoI*, *AvrII* and *XhoI* restriction sites. Subsequently, the CYP was cloned via InFusion into the *ATR2tr*-pET28 construct.

The primers employed in the process are listed in Table 4.1. Plasmids containing the codon optimised *ATR2* and *CYP73A5* sequences were retrieved from the Bruce's group stock.

**Table 4.1:** Oligonucleotides used for the PCR amplification of the DNA sequences encoding the *ATR2* reductase and the S-CYP73A5tr.

Protein	Primer sequence
<b>ATR2</b>	Forward - 5' TCGGCCGCACTCGAGAGCGGTCTAGGGGTAGCGGTAATAGCAAAC 3'
	Reverse - 5' GGTGGTGGTGCTCGAGCCAAACATCACGCAGATAACGACCGCTGGT 3'
<b>CYP</b>	Forward - 5' AGGAGATATACCATGGATGGCCAAAAAACCAGCAGCA 3'
	Reverse - 5' GCTGGTGCTACCGGATCCTGAGTTACGCGGTTTCAT 3'

The mix for the PCR amplification reactions consisted of 10 mM dNTPs mix, 20  $\mu$ M of each primer, 50 ng of template DNA and 1.25 U of *Pfu* DNA polymerase in 10x *Pfu* buffer and water to a final volume of 50  $\mu$ l. The cycles included a first denaturation step at 95 °C for 2 min, then 95 °C for 30 s, an annealing step at 50 °C for 30 s and extension at 74 °C for 2 min for a total of 30 cycles and a final extension step at 74 °C for 10 min. For the InFusion cloning procedure see Materials and Methods, paragraph 2.4.6. The final construct consisted of the N-terminal 'S' solubility tag (sequence: MAKKTSSKG, for clarification see Chapter 3), followed by the truncated CYP enzyme, directly bound through a poly-glycine/serine linker to the truncated *ATR2* reductase, with a C-terminal hexa-histidine tag (figure 4.3). The vector was transformed into competent *E. coli* cells (with the heat-shock procedure, see materials and methods paragraph 2.4.8), harvested from the cultures via plasmid prep (as described in materials and methods, paragraph 2.4.5), sent for sequence verification and used to initiate subsequent cultures for protein expression studies.



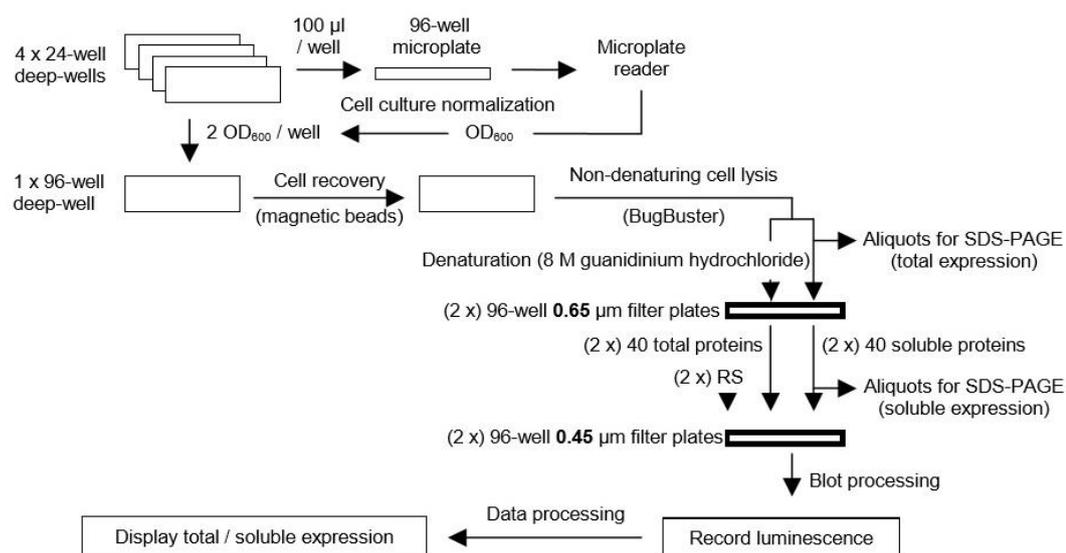
**Figure 4.3:** Map of the fusion platform. The cloning site is highlighted in blue. The fusion construct consists of an N-terminal S-CYPtr domain directly connected through a poly-serine/glycine linker to the C-terminal histidine-tagged ATR2tr redox partner (S-tag = MAKKTSSKG, linker: GSTSSGSG).

### 4.2.2 First expression trials

The first expression trial for the pAtHIA2 construct, containing CYP73A5tr, was performed using *E. coli* Arctic express as expression host, following the procedure which allowed the soluble expression of Arabidopsis P450s S-CYP73A5tr, S-CYP81D11tr and S-CYP81A9tr (see details in Chapter 3, section 3.3.2). Briefly, the transformed cells were inoculated in LB medium and then subcultured in M9 medium. Protein expression was induced with 0.5 mM IPTG and cultures were harvested after 20 h of incubation at 16 °C, shaking at 180 rpm.

### 4.2.3 High-throughput automated expression screening

A team, led by Dr. Jared Cartwright at the Protein Production Technology Facility of the University of York, developed an automated high-throughput expression screening platform, based on the system first described by Vincentelli *et al.* (201). Dr. Cartwright's system involved the use of a Freedom EVO robot (Tecan), guided by a specifically-designed macro program, which allowed the setup and screening of different parallel expression conditions, on a 48-well plate.



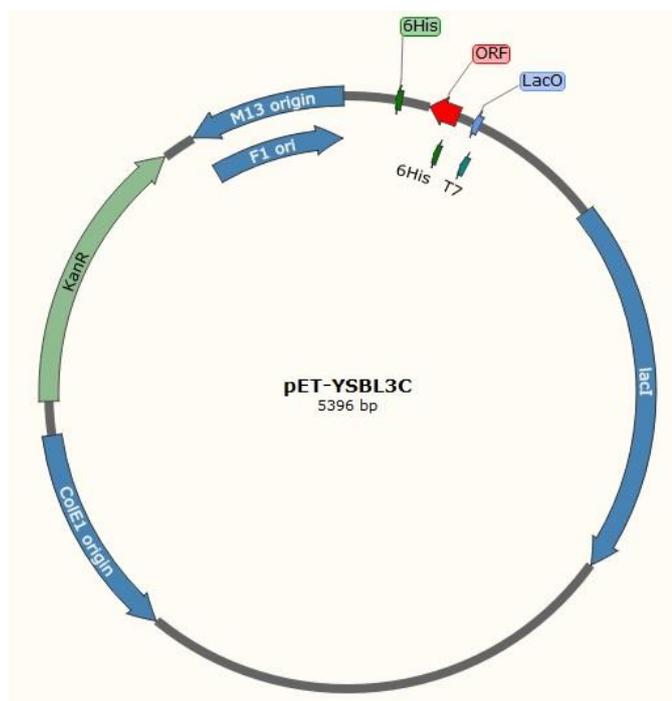
**Figure 4.4:** Overview of the automated process for protein expression screening designed by Vincentelli *et al.* (201). Cell cultures, in 2-ml format, are set up in multi-well microplates, allowing a quick spectrophotometric evaluation of cell density ( $OD_{600}$ ) using the microplate reader. Cells are harvested via centrifugation and disrupted chemically. A round of centrifugation allows the separation of the soluble fractions from the cell debris and collected samples (from the total lysate and from the soluble fraction) are blotted onto nitrocellulose for western blot analysis. All the steps are controlled through a computer and performed by a liquid-handling robot.

The fusion construct of CYP73A5tr-ATR2tr (without the 'S' solubility tag) was re-cloned via InFusion into five different expression vectors, named pETFPP 1 to 5. The oligonucleotides used for the PCR amplification of the whole CYP73A5tr-ATR2tr fused construct, as well as of the single CYP73A5tr and ATR2tr domain (each cloned and expressed alongside the fusion, as controls) are listed in Table 4.2.

**Table 4.2:** List of the oligonucleotides used as primers for the introduction of the S-CYPtr, ATR2tr and the CYP73A5tr-L-ATR2tr fusion (using the 'F-CYP73A5tr' and the 'R-ATR2tr' primers for the CYP/ATR fusion construct) in the expression vectors.

Oligo name	Primer sequence
<b>F-CYP73A5tr</b>	5' – TCCAGGGACCAGCAATGGCCCTGCGTGGTAAAAAACTG – 3'
<b>R-CYP73A5tr</b>	5' – TGAGGAGAAGGCGCGTTATGAGTTACGCGGTTTCATCAC – 3'
<b>F-ATR2tr</b>	5' – TCCAGGGACCAGCAATGGGTAGCGGTAATAGCAAACG – 3'
<b>R-ATR2tr</b>	5' – TGAGGAGAAGGCGCGTTACCAAACATCACGCAGATAACG – 3'

The pETFPP\_1-5 expression vectors used in this automated screening all presented a pET-YSBL3C backbone (reported in figure 4.5), with a cleavable N-terminal hexa-histidine tag, for detection purposes, followed by one of four solubility tags (listed in Table 4.3). The tags included in the new open reading frame were the widely used maltose-binding protein (MBP (202-205)), glutathione transferase (GST (203, 206, 207)), colicine immunity protein 9 (Im9, NCBI Reference Sequence: YP\_002533538.1 (208)), and green fluorescent protein (GFP (209, 210)).



**Figure 4.5:** Map of the pET-YSBL3C vector, used as a backbone for the pETFPP\_1-5 constructs. This expression vector is characterised by a T7 promoter, an N-terminal hexa-histidine tag, an HRV-3C cleavage site and the *Bse*RI, *As*cl, *Nde*I restriction sites for insert cloning. In the pETFPP\_1-5 expression vectors, an additional N-terminal tag is present, downstream of the his-tag.

**Table 4.3:** Structure of the pETFPP\_1-5 expression vectors developed by Dr. Cartwright (HIS = hexa-histidine tag, MBP = maltose-binding protein, GST = glutathione transferase, Im9 = colicine immunity peptide 9, GFP = green fluorescent protein).

Construct name	N-terminus structure
pETFPP_1	HIS-3CProtease-ORF
pETFPP_2	HIS-MBP-3CProtease-ORF
pETFPP_3	HIS-GST-3CProtease-ORF
pETFPP_4	HIS-Im9-3CProtease-ORF
pETFPP_5	HIS-GFP-3CProtease-ORF

In the first screening, *E. coli* BL21(DE3) was chosen as host for expression and two different growth media, LB and AI (see composition in Materials and Methods, section 2.3), were tested. After an overnight incubation of the BL21(DE3) transformants at 37 °C, single colonies were picked from the selection plates and inoculated in each of the 2-ml wells of the screening platform, containing either LB or AI medium. The bacterial cultures were grown at 37 °C at a shaking speed of 660 rpm, until induction with 1 mM IPTG (only for the LB cultures), then the temperature was set to 16 °C for protein expression. After 24 h, the cell density (OD<sub>600</sub>) of all the cultures was measured and each adjusted to the minimum value recorded, for normalisation. The expression plate was subjected to centrifugation at 6,400 rpm to harvest the cultures. The cells were then disrupted chemically with a mixture of lysozyme and BugBuster Protein Extraction Reagent (both from Novagen) and the soluble fractions were separated from the cell debris via centrifugation. Following a treatment with 5 M guanidinium hydrochloride to denature the proteins, the fractions were blotted on a membrane and subsequently incubated with HRP conjugated antibodies (from Qiagen), raised against the his-tags. In this way, the result of the expression screening could be visualised as a dot plot, whereby the intensity of each spot (corresponding to each blotted culture/fraction) was proportional to the level of expression of the his-tagged protein. The detected product could be quantified by comparing the luminosity of each dot to the luminosity of his-tagged standards.

#### 4.2.4 Purification of the fused construct

For the complete purification procedure see Materials and Methods, paragraph 2.5.6.

Constructs yielding the highest levels of expression detected in the screen were identified for further study. The expression format was scaled-up from 2 ml to 500 ml cultures. The cells were harvested via centrifugation, resuspended in phosphate buffer supplemented with 200  $\mu$ M PMSF and 250-500 mM NaCl and lysed with the sonicator. The soluble fraction was separated from the cell debris via centrifugation and filtered prior to the loading on the Ni-affinity chromatographic column. The result of the purification was verified with an SDS-PAGE gel of the fractions collected throughout the process.

#### 4.2.5 Activity assay of the fused construct CYP73A5-ATR2

The activity of the fusion protein was assayed using whole *E. coli* cells. After 24 h of growth at 16 °C the cells were harvested via centrifugation and resuspended in 50 mM potassium phosphate buffer, pH 7 (100 mg cells / ml of buffer). The suspension of whole cells was used to assess the ability of the construct to catalyse the hydroxylation of cinnamic acid (Sigma-Aldrich), the natural substrate of AtCYP73A5, to *p*-coumaric acid (Sigma-Aldrich). The reaction, carried out at 28 °C and 200 rpm, was initiated by the addition of 200  $\mu$ M cinnamic acid to the cell suspensions (2 ml each, in 50 ml falcon tubes). Parallel reactions with cells transformed with the empty vector were carried out as negative controls. Samples were collected over a period of seven h. The reactions were terminated with an equivalent volume of methanol and 35  $\mu$ l of the supernatant after centrifugation were analysed using an HPLC (Waters Alliance 2695 separation module equipped with a Waters 2996 photodiode array detector) with a Waters X-Bridge C18 column heated to 30 °C, employing a gradient of water (A) and methanol with 0.1% acetic acid (B): 90 % A for the first 3 min, 90 to 55 % A in 1 min, 55 % B for A min and 90 % A for 5 min (separation method designed by Dr. Schüchel (211) , modified from Chen and Morgan (212)). With this separation method, cinnamic acid and *p*-coumaric acid eluted at 11.7 and 8.1 min, respectively.

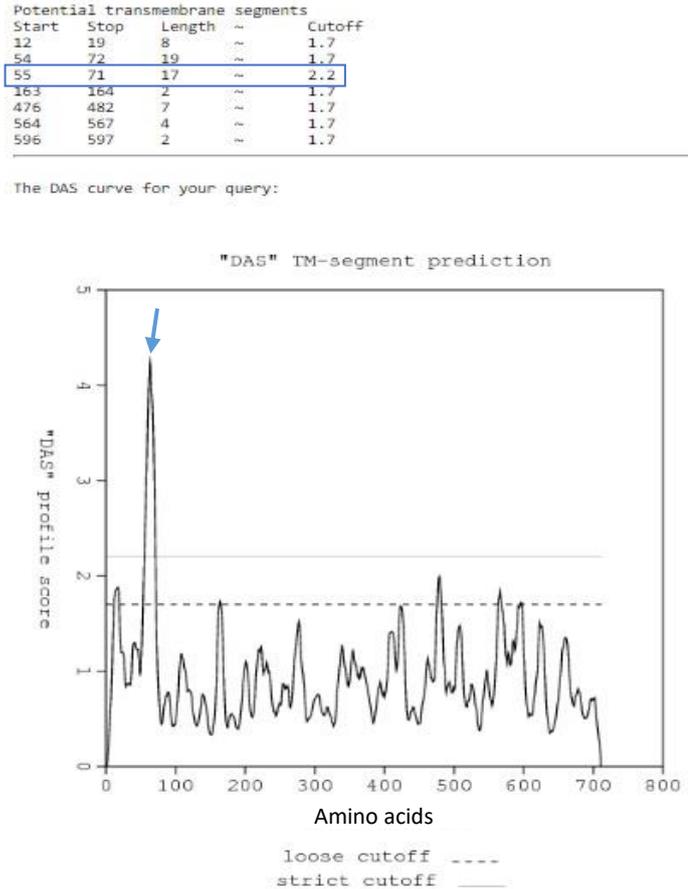
## **4.3 Results**

### **4.3.1 Creation of the pAtHIA2 fusion platform**

The pAtHIA2 fusion platform was created starting from the production of the ATR2tr-pET28 construct, followed by the introduction, upstream of the reductase, of the S-CYPtr with Infusion-compatible ends.

The sequence analysis of the ATR2 reductase, carried out using the Das and TMHMM online transmembrane structure prediction tools, revealed that the first 72 amino acids were likely to form the membrane-binding domain of the protein (figure 4.6).

(A)

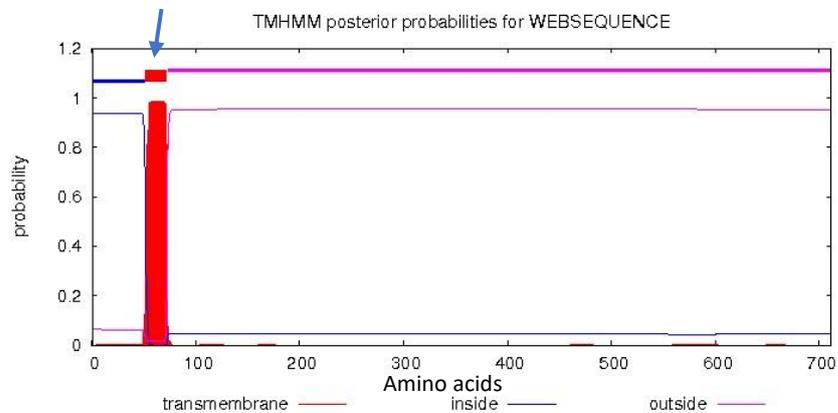


(B)

### TMHMM result

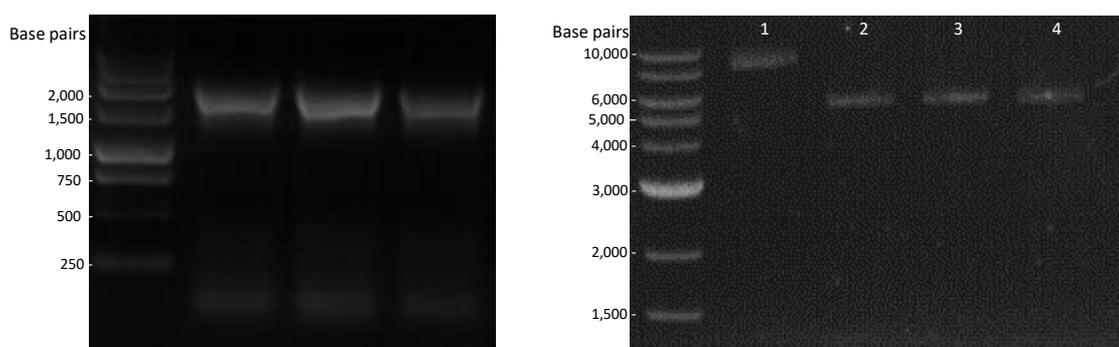
[HELP](#) with output formats

```
# WEBSEQUENCE Length: 711
# WEBSEQUENCE Number of predicted TMHs: 1
# WEBSEQUENCE Exp number of AAs in TMHs: 20.51299
# WEBSEQUENCE Exp number, first 60 AAs: 8.44722
# WEBSEQUENCE Total prob of N-in: 0.93533
WEBSEQUENCE TMHMM2.0 inside 1 51
WEBSEQUENCE TMHMM2.0 TMhelix 52 72
WEBSEQUENCE TMHMM2.0 outside 73 711
```



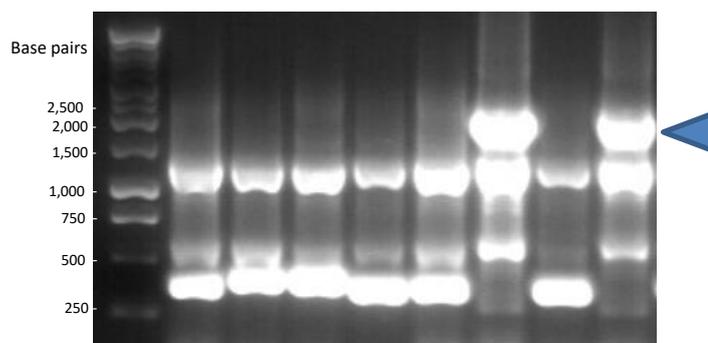
**Figure 4.6:** Transmembrane domain predictions by DAS (panel A) and TMHMM (panel B) for *AtATR2*. As highlighted in blue on the results panels and with an arrow on the corresponding graphs, the algorithms identified the region between amino acids 55-71 and 52-72 as the one forming the N-terminal transmembrane region.

For this reason, the oligonucleotides for the PCR amplification of the ATR2 sequence were specifically designed in order to allow the extension of a truncated version of the reductase (ATR2tr), while also introducing the *Nco*I, *Avr*II and *Xho*I flanking restriction sites. In parallel, the pET28 expression vector was double-digested with the corresponding *Nco*I and *Xho*I restriction endonucleases, following the instructions from the manufacturer. The outcome of the PCR amplification and of the digestion are reported in figure 4.7.



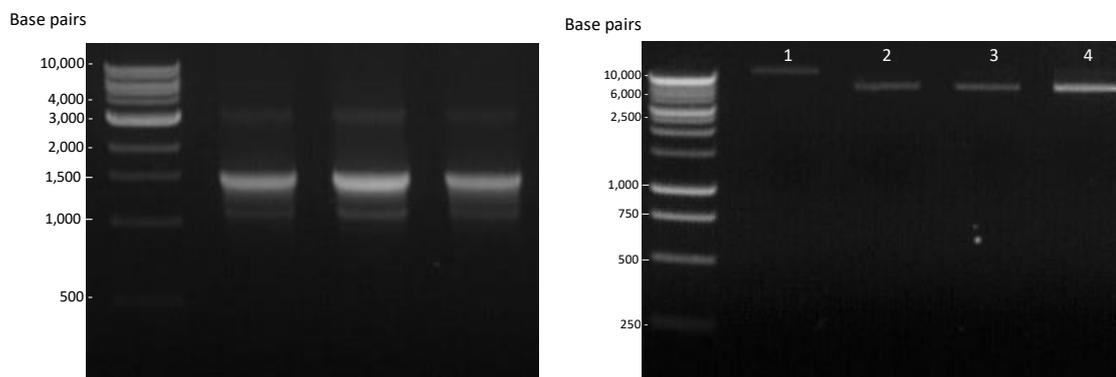
**Figure 4.7:** Agarose gel analysis of the PCR amplification of ATR2tr (left panel, in triplicate; expected size of the amplicon: ~1.9 Kbases) and the double digestion of the pET28 vector with *Nco*I and *Xho*I (right panel). 1: undigested plasmid, 2: *Nco*I single cut, 3: *Xho*I single cut, 4: double cut.

The ATR2tr amplicons were inserted using InFusion in the *Nco*I/*Xho*I-linearised pET28 vector and the cloning reaction was transformed into competent *E. coli* cells, which were plated onto selective LB agar plates. The outcome of the PCR amplification (using primers specific to the T7 promoter and T7 terminator regions) performed on the colonies picked from the selective plates is shown in figure 4.8.



**Figure 4.8:** Colony PCR for the verification of the presence of ATR2tr in the newly-built construct. The blue arrow indicates the positive colonies, presenting DNA bands with the correct size. Expected size of the amplicon: ~1.9 Kbases.

The *S-CYP73A5tr* encoding sequence was amplified via PCR, to introduce the *NcoI* and *AvrII* flanking restriction sites. In parallel, the pET28-ATR2tr plasmid was double-digested with the corresponding endonucleases (figure 4.9).



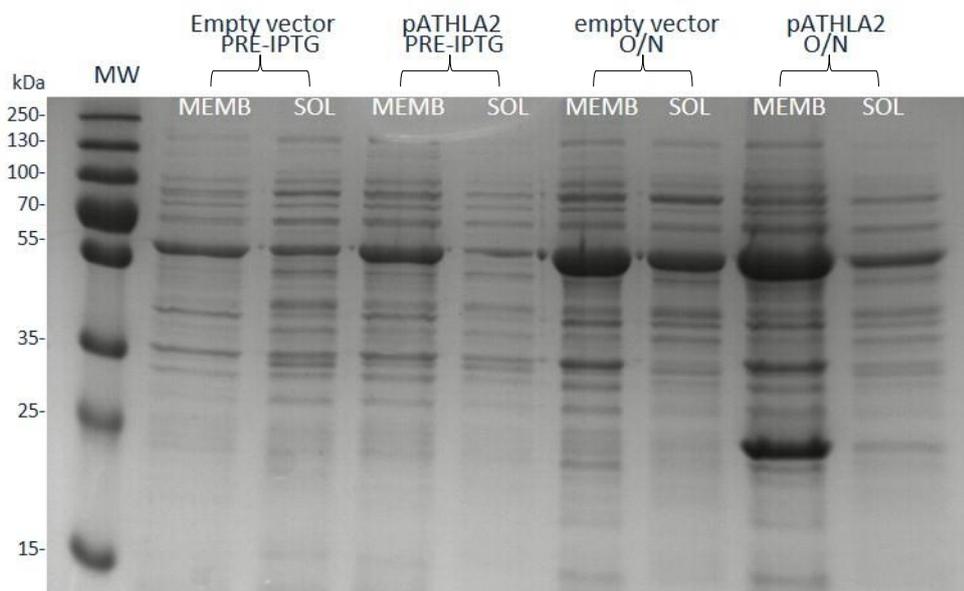
**Figure 4.9:** Agarose gel analysis of the amplification of the *S-CYP73A5tr* insert (left panel, in triplicate; expected size of the amplicon: ~1.5 Kbases) and of the double digestion of the pET28-ATR2tr plasmid with *NcoI* and *AvrII* (right panel). 1 = undigested plasmid, 2 = *NcoI* single cut, 3= *AvrII* single cut, 4 = double *NcoI/AvrII* cut).

The *S-CYP73A5tr* amplicon and the linearised pET28-ATR2 plasmid were purified and mixed for the InFusion cloning reaction. The mix was directly transformed into competent *E. coli* cells for subsequent purification and sequence verification.

### 4.3.2 First expression trial of the pAthIA2 (*S-CYP73A5tr*-ATR2tr) fusion in *E. coli* Arctic express

The first expression trial of the newly-built pAthIA2 fusion was performed using the same conditions as for the expression of the *S-CYPtr* constructs in Chapter 3.

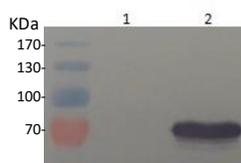
The fusion construct was transformed into Arctic express *E. coli* cells and the M9 cultures were harvested after 16 h from the IPTG induction. The result of the expression is reported in figure 4.10.



**Figure 4.10:** SDS-PAGE analysis of the outcome of the expression trial of pATHLA2 (the fusion of S-CYP73A5tr – ATR2tr), employing the A. express/M9 expression procedure developed in Chapter 3. PRE-IPTG = before induction, O/N = overnight, MEMB = membrane fraction, SOL = soluble fraction. Expected size of the product: ~127 kDa.

A band at the expected size of approximately 127 kDa was not observed on the SDS-PAGE analysis of the samples collected after the expression induction for the Arctic express cultures (alongside the ones transformed with the empty vector, used as negative controls). In order to verify if this result was due to the presence of the reductase domain, an expression trial was performed with a construct consisting of the ATR2tr reductase cloned into a pET-YSBL3C vector, without the P450 domain.

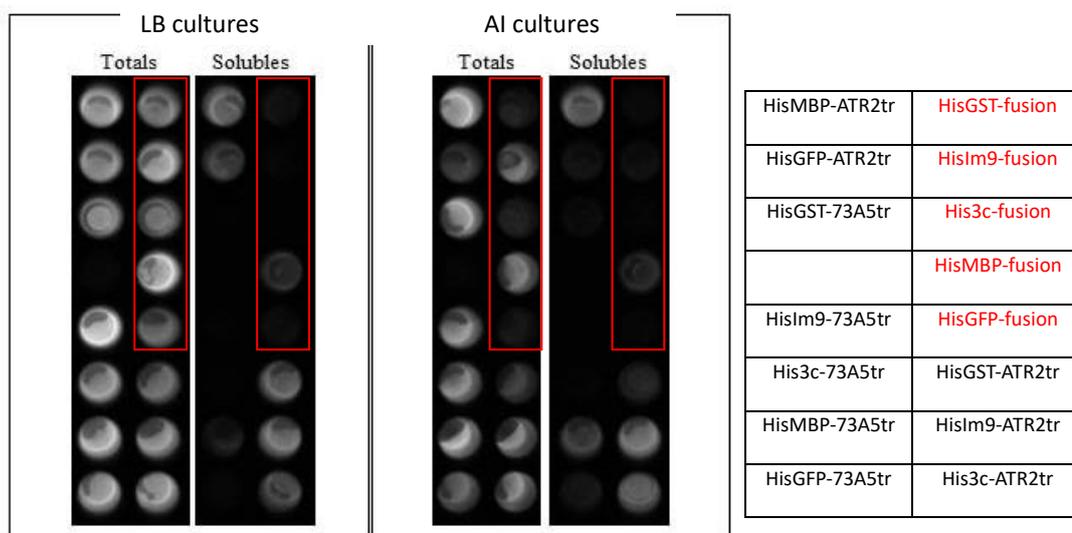
As shown in figure 4.11, the N-terminally truncated ATR2tr was expressed successfully using the Arctic express/M9 protocol.



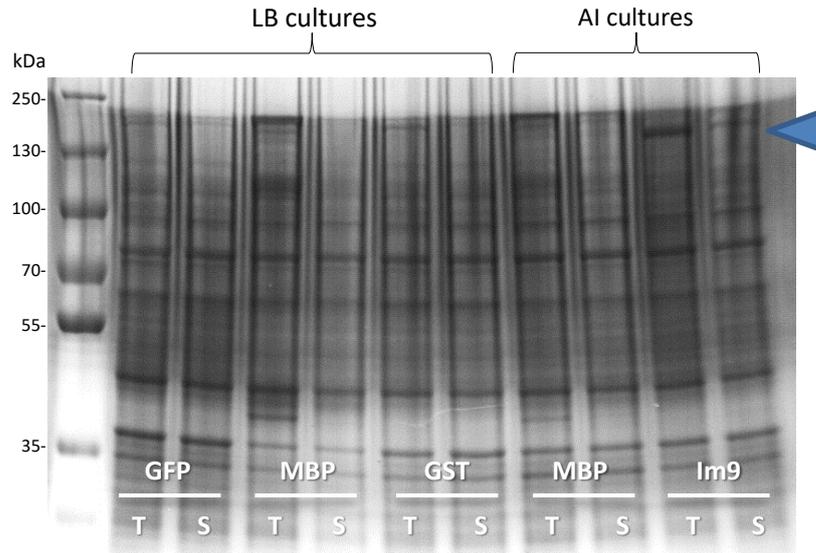
**Figure 4.11:** Western blot analysis of the expression of ATR2tr cloned in pETYSBL-3C in *E. coli* Arctic express. 1 = empty vector control, 2= ATR2tr in the soluble fraction. Expected molecular weight of the product: 71.5 kDa.

### 4.3.3 Re-cloning and automated protein expression screening

The fusion of the Arabidopsis N-terminally truncated CYP73A5, connected with a polyGlySerThr linker (GSTSSGSG) to the truncated ATR2 reductase was cloned into five different pET-YSBL3C-based vectors (pETFP\_1-5), presenting five alternative N-terminal tags. At the end of the expression trial, carried out using BL21(DE3) *E. coli* as host and AI and LB as growth media, the His-tagged products were detected through western blot (figure 4.12) and the molecular weight of the positive hits (MBP-AI/LB, GST-LB, GFP-LB, Im9-AI, within the red boxes) was estimated via SDS-PAGE (figure 4.13).



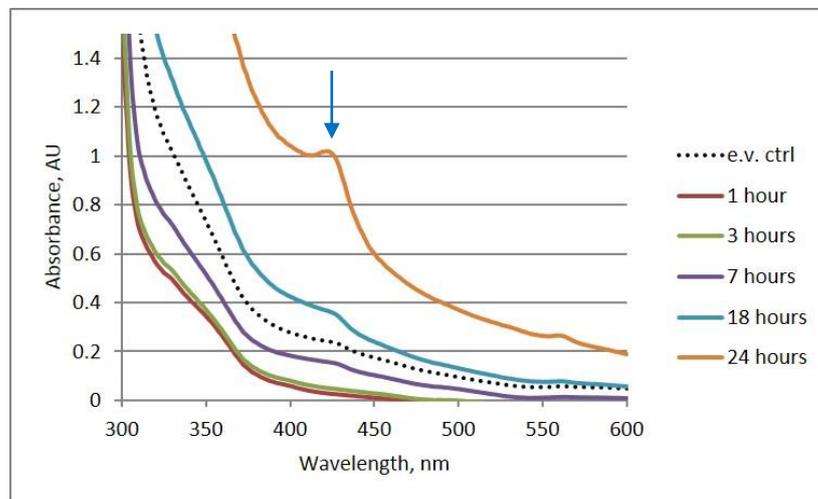
**Figure 4.12:** Western blot analysis of the membrane blotted with the fractions from the expression trial. Each dot corresponds to a specific construct. His-tagged products in the total lysate and in the soluble fraction of the cells grown in LB (left block) and AI (right block) media are detected by using antibodies anti his-tag. The scheme on the right illustrates the position of the constructs in each of the four panels. Highlighted in the red boxes are the fusion constructs.



**Figure 4.13:** SDS-PAGE analysis of the his-tagged products detected in the previous western blot of the BL21(DE3) cultures. GFP = construct with GFP tag; MBP = construct with MBP tag; GST: construct with GST tag; Im9 = construct with Im9 tag. For each culture, the total lysate (T) and the soluble fraction (S) were loaded in parallel.

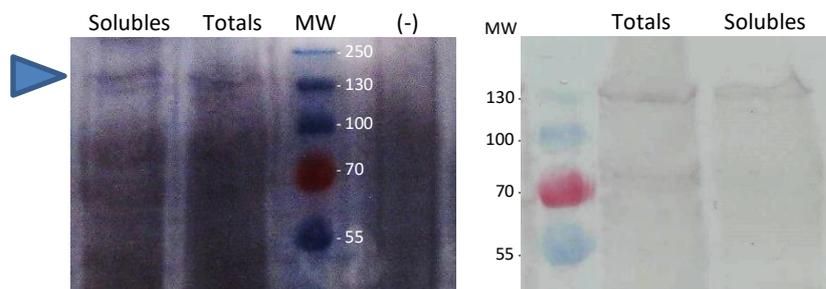
The SDS-PAGE gel (figure 4.13) showed a promising product band with the correct size for the Im9 construct, expressed in AI medium.

Given the promising result, the expression of the Im9 fusion (Im9-CYP73A5tr-L-ATR2tr) was scaled up from 2 ml to 500 ml. The production of the Im9 fusion was monitored spectrophotometrically, by recording the absorbance spectrum between 300 nm and 600 nm, at different time points post induction (figure 4.14).



**Figure 4.14:** Overlap of the spectra recorded for the soluble fractions of the cell lysates, over a period of 24 h after the induction. The spectrum for the negative control (e.v. ctrl = empty vector) was recorded at the final time point.

A clear peak at around 420 nm could be observed for the sample collected after 24 h from the induction time. Further evidence for expression was obtained by analysing the size of the proteins on an SDS-PAGE gel and via western blot (figure 4.15).

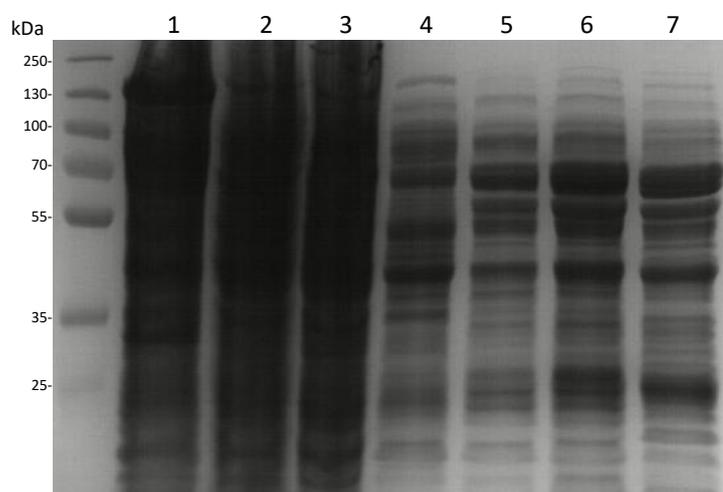


**Figure 4.15:** SDS-PAGE and western blot analysis of the Im9-CYP73A5 fusion expression scale-up. Expected molecular weight of the product: 137 kDa. The (-)CTRL corresponds to the cells expressing the empty vector.

The results in figure 4.15 show that the Im9-CYP73A5tr-ATR2tr fusion was expressed in the total lysate and the soluble fractions of the cells.

#### 4.3.4 Purification of the Im9-CYP73A5 fusion construct

The separated soluble fractions of lysed BL21(DE3) cells expressing the construct were subjected to nickel-affinity chromatography for isolation. The samples collected throughout the process were loaded on an SDS-PAGE gel to verify the presence and purity of the desired product (figure 4.16).



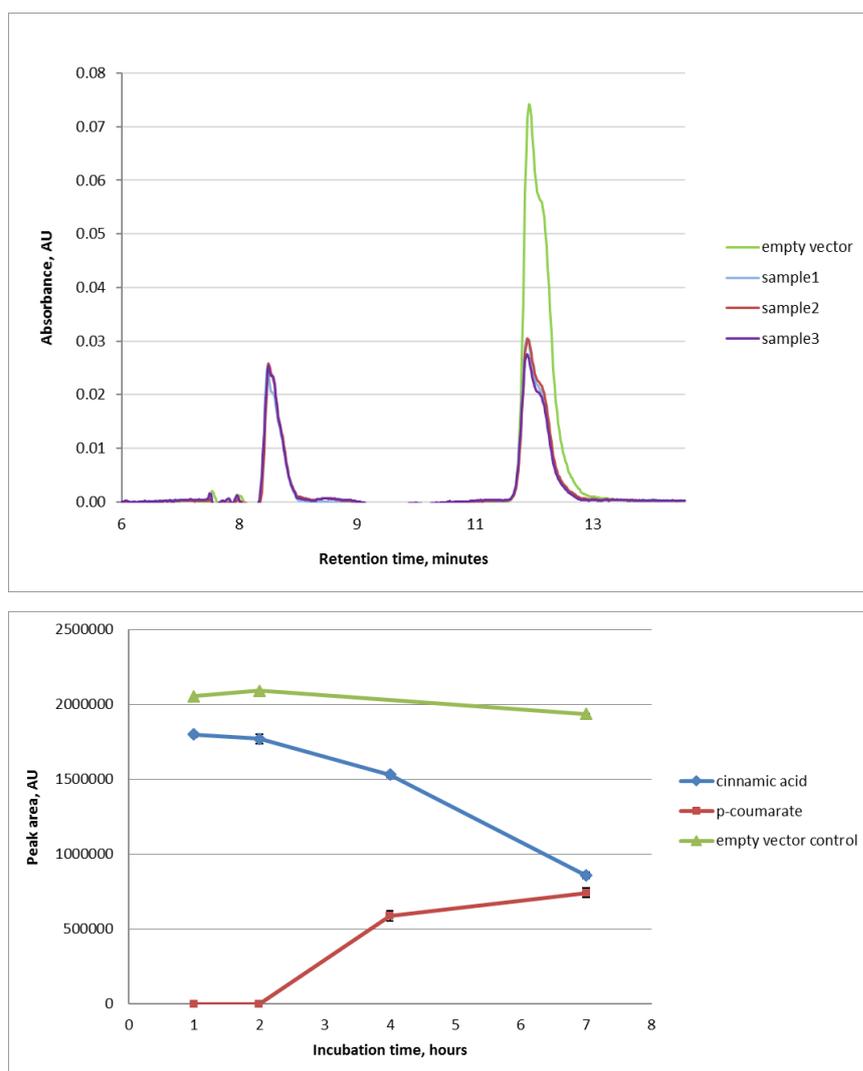
**Figure 4.16:** SDS-PAGE analysis of the fractions collected during the purification process of the Im9-CYP73A5tr-ATR2tr fusion protein (1= total lysate, 2 = soluble fraction, 3= unbound, 4= wash 1, 5-7= elutions).

The 137 kDa fusion protein was not present in the eluates. A UV-Vis spectrophotometric analysis of the 'unbound' fraction (well 3, in figure 4.16) showed that the expected protein did not bind efficiently to the nickel resin. Further purification attempts were carried out, by diluting the cell suspensions (7.5 ml of lysis buffer / g of cells) or using different buffers (50 mM potassium phosphate pH 8 with 500 mM NaCl, 50 mM sodium phosphate pH 8 with 300 mM NaCl, or 50 mM Tris-HCl pH 8 with 250 mM NaCl) or employing a batch purification process, where the protein extract was incubated for a long period (overnight) with HIS-Select Nickel Affinity Gel (Sigma Aldrich). In all cases, the full length fusion product was not obtained.

#### **4.3.5 Activity assay of the fused Im9-CYP73A5tr-ATR2tr construct**

As purified Im9-CYP73A5tr-ATR2tr protein could not be obtained, the activity assays to test the Im9 fusion protein were carried out on whole cells transformed with the construct.

Cell suspensions in phosphate buffer at pH 7, supplemented with 100  $\mu$ M of cinnamic acid, the natural substrate of AtCYP73A5, were monitored over a period of 7 h. The resulting chromatograms at the end of the incubation and the conversion of cinnamic acid into *p*-coumaric acid are shown in figure 4.17.



**Figure 4.17:** HPLC analysis of the samples after 7 h of incubation of BL21(DE3) whole cells expressing the fusion Im9-CYP73A5tr-ATR2tr construct with 100  $\mu$ M cinnamic acid. Top panel: chromatogram showing the formation of the product, *p*-coumaric acid, in comparison to the negative control (empty vector) at the end of the incubation period. Bottom panel: conversion of cinnamic acid (blue diamonds) into *p*-coumaric acid (red squares) over time; for the empty vector (-) ctrl (green triangles) only the concentration of cinnamic acid is reported. Assay performed in triplicate  $\pm$  SE.

The assay showed that the expressed Im9-fusion was active in whole cells, with the conversion of more than 59 % of cinnamic acid to *p*-coumaric acid (final yield:  $37.1 \pm 1.6 \mu$ M) over a period of six h.

## 4.4 Discussion

Cytochromes P450 enzymes are defined as NAD(P)H-dependent monooxygenases, strictly dependent on the presence of a co-expressed and co-localised redox partner. Exceptions to the general rule are some bacterial and fungal P450 macroproteins, characterised by a P450 heme domain structurally connected to the redox partner, forming a highly active ~110-120 kDa polypeptide. Important examples of these P450 fusions are the widely studied *B. megaterium* P450 BM-3 (with multiple homologues found also in other *Bacillus* species (213, 214)), the *Rhodococcus* sp. strain NCIMB 9784 P450Rhf (89), and the *F. oxysporum* P450foxy (215).

Besides the STORR ((S)-to-(R) reticuline) protein discovered in 2015 by Winzer *et al.* in opium poppy (*Papaver somniferum* L.), no other natural plant P450 fusion has been described so far. The STORR multidomain macroprotein was identified during the discovery of the pathway of enzymes involved in the synthesis of morphinan alkaloids in opium poppy. Specifically, STORR consists of a P450 heme domain (CYP82Y2) connected at its carboxy terminus to an oxidoreductase domain. Comparing the phenotypes of different mutants at the specific *storr* locus, the researchers discovered that, interestingly, the P450 and the oxidoreductase catalyse two different sequential steps of the epimerisation of (S)-reticuline to (R)-reticuline, one of the intermediates in the synthesis of morphine and codeine. This meant that the C-terminal domain of the fusion was not a P450 redox partner, but instead, it was a protein with a specific catalytic function, with or without the P450-heme domain (216).

With the final goal of creating a drop-in construct for the expression and functional characterisation of redox self-sufficient orphan plant CYPs, the main objective of this thesis chapter was the creation of an active plant P450-P450reductase fusion platform. With this purpose, a number of constructs were generated and screened. Starting from the results achieved by Dr. Schuckel, the *AtATR2* reductase and the poly-glycine/serine linker were kept in the new artificial constructs; *AtCYP73A5* was used to validate the activity of the chimera. Both the sequences of the reductase and the P450 were codon optimised for *E. coli* expression. Additionally, these two protein

sequences were truncated of the N-terminal membrane-binding domain, identified with the Das and TMHMM online prediction tools, to increase the chances of having an overall soluble product. The 'S' peptide (sequence: MAKKTSSKG, see Chapter 3) as well as the same expression host (*E. coli* Arctic express) and conditions that enabled the expression of three different plant P450s in Chapter 3, did not yield the expected 127 kDa fusion polypeptide. It may be hypothesised that the bulky pAtHIA2 construct (S-CYP73A5tr-L-ATR2tr) could not be efficiently encoded by the translational machinery in the chosen expression host, which, alongside the exogenous proteins also co-expresses chaperones. Requiring a quicker and more efficient method to screen multiple conditions at once, an automated liquid-dispensing robot system was employed. The screen identified that the colicin immunity protein 9, Im9, effectively helped in the soluble expression of the whole P450/P450-reductase construct. The 86-amino acid Im9 protein is encoded by *E. coli* cells as a self-defence system from the cytotoxic effects of colicins, which are compounds naturally produced against other competing bacteria. Specifically, Im9 binds tightly to the E9 colicin and dissociates from it once the complex reaches specific receptors on 'rival' cells (217). This solubilising effect on the CYP73A5tr-L-ATR2tr fusion could be due to the acidity of Im9 (pI: 4.2-4.5, calculated net charge at pH 7: - 8.8), as reported for other acid solubility peptides used in protein expression screenings by Su *et al.* (218). Alternatively, or perhaps additionally, Im9 could interact directly with the target proteins, minimising the chance of protein aggregation.

Several attempts were made to purify the Im9-CYP73A5 fusion protein, using a range of conditions, but with no success. While more research is required to determine the optimal conditions for the purification of the soluble Im9-CYP chimera, a number of hypotheses could be drawn to explain why it was not successful this time. First of all, it could be possible that the N-terminal histidine tag was degraded or not readily accessible and, for this reason, the protein could not bind to the column. This was demonstrated by the fact that the product of the expected size of 137 kDa could be found, as shown in the SDS-PAGE gel verification, in the first washing fractions, during and right after the sample loading onto the chromatographic column. In addition, another factor to take into consideration is the status of the column. It might be

possible that the resin was not properly regenerated nor homogenously equilibrated, with resulting low binding capacity. Proposed approaches for new purification trials would be the re-positioning of the histidine tag at the C-terminus, the use of different binding buffers or the switch to a different column, such as for example the 2'5'-ADP Sepharose column. In this particular case, the resin, consists of sepharose beads coated with 2',5'-ADP (adenosine 2',5'-diphosphate, a structural analogue of NADP), which binds tightly to molecules that require NADP as cofactor (thus presenting NADP binding sites), without the need of an affinity tag. This approach has been successfully employed by Wadsäter *et al.* for the purification of the P450 reductase from *S. bicolor* (219). Given the lack of purified Im9-CYP73A5 fusion protein, to determine the integrity of the construct, activity assays with the natural substrate of CYP73A5 were carried out using whole cells expressing the construct. Cells incubated with cinnamic acid were able, as compared to cells transformed only with the empty vector (negative controls), to convert almost 60 % of the substrate, generating more than 30  $\mu$ M of the hydroxylated product, *p*-coumaric acid, in a time frame of six h. Longer assays were also performed, but the chromatograms were not as clear as for the previous time points, possibly due to the formation of other unspecific compounds by the *E. coli* cells.

Considering these promising results, the Im9 construct, with a C-terminal *AtATR2tr* reductase could be used for the cloning, expression and characterisation of a plethora of plant P450s that are still uncharacterised to date. The cloning procedure, consisting of a single digestion and InFusion reaction is relatively easy and the coupling with an automated liquid sampling robot could significantly decrease the time and resource consumption related to the optimisation of the other expression parameters, such as expression host, growth media and temperatures.

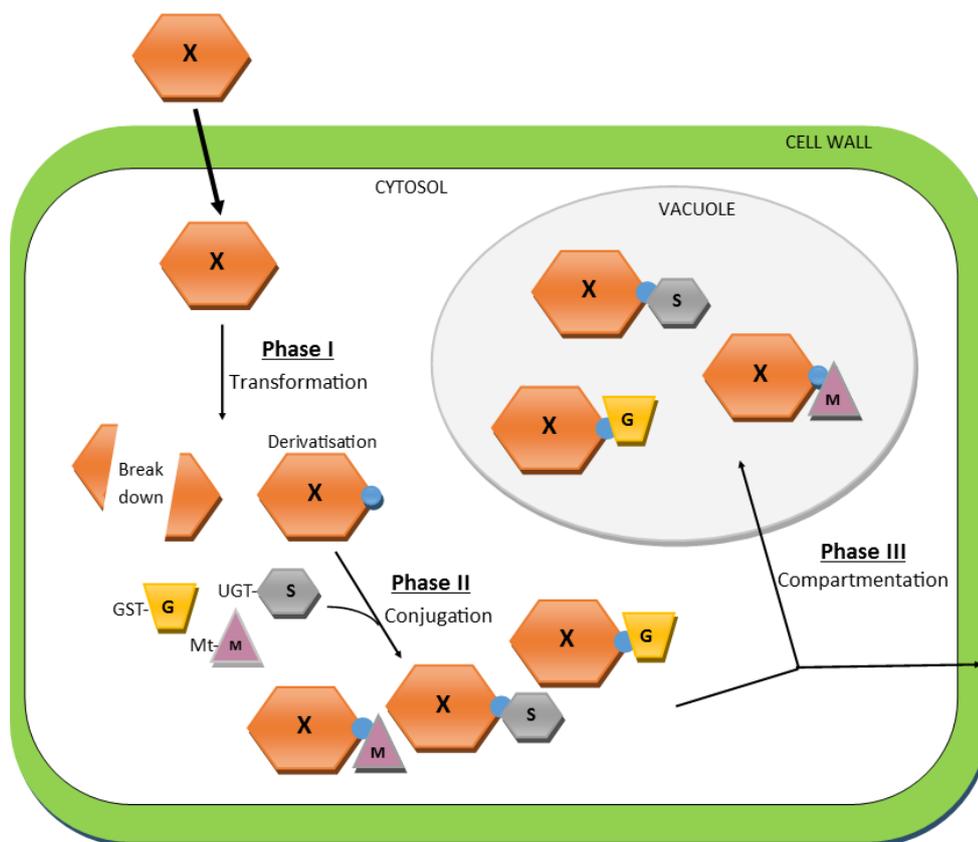
## **Chapter 5: A role for CYP81D11 in the detoxification of the explosive TNT**

### **5.1 Introduction on the role of Cytochromes P450 in plant detoxification processes**

Plants exhibit an extensive collection of CYP genes, with more than 13,000 sequences currently annotated (220). These CYPs are key in numerous pathways related to structural development (lignins), defence mechanisms and signaling cascades (oxylipins, phytoalexins, glucosinolates, alkaloids, cyanogenic glucosides) and biosynthesis of hormones (gibberellins, brassinosteroids, abscisic acid, cytokinins) (15, 17, 221). Furthermore, CYPs are also involved in the synthesis of secondary metabolites of chemical/pharmaceutical value, such as artemisinic acid (used in antimalarial combination therapies (222, 223)), the anticancer compound taxol (paclitaxel, (112)), fragrances, flavours and nutraceuticals. Moreover, plant CYPs can also be responsible for the breakdown of endogenous molecules as well as of xenobiotics (221, 224).

Similar to hepatic CYPs in mammals, plant CYPs are able to detoxify - mainly via hydroxylation, dealkylation and oxidation - xenobiotic compounds. Among these are organic pollutants, such as herbicides and pesticides, taken up from the surrounding environment (224). The first evidence for P450-mediated xenobiotic detoxification was collected by Frear and coworkers almost 50 years ago: they observed that an oxidase from cotton microsomal fractions was able to dealkylate monuron (3-(4'-chlorophenyl)-1,1-dimethylurea), a phenylurea herbicide. The connection to CYPs was easily made, considering the location of the active enzyme and the requirement of molecular oxygen as well as of NADPH (225, 226). Later, Mougín *et al.* confirmed that the herbicide chlortoluron (3-(3-Chloro-4-methylphenyl)-1,1-dimethylurea) was metabolised via dealkylation and hydroxylation by a set of microsomal CYPs in wheat (*Triticum aestivum*) (227). Following from Fonne-Pfister and colleagues' studies on

xenobiotic metabolism in Jerusalem artichoke (*H. tuberosus*) (228), Robineau *et al.* determined that CYP76B1 was able to catalyse the dealkylation of further phenylurea herbicides, such as chlortoluron and isoproturon (180, 229). In all the examined cases, the result of CYP catalysis was the conversion of the compounds to more hydrophilic and less toxic derivatives, therefore contributing directly to the overall plant tolerance and resilience to such xenobiotics. All these early observations opened up the doors to the study and application of CYP enzymes to agronomics and bioremediation. Besides cytochromes P450, the metabolism of agrochemicals and pollutants in plants is also carried out by other classes of enzymes, including UGTs, GSTs and malonyl transferases, which are responsible for the conjugation of the compounds to molecules such as sugars, glutathione or malonate, respectively (230). These bio-converted compounds are then transferred to vacuoles or cell walls for storage (figure 5.1).

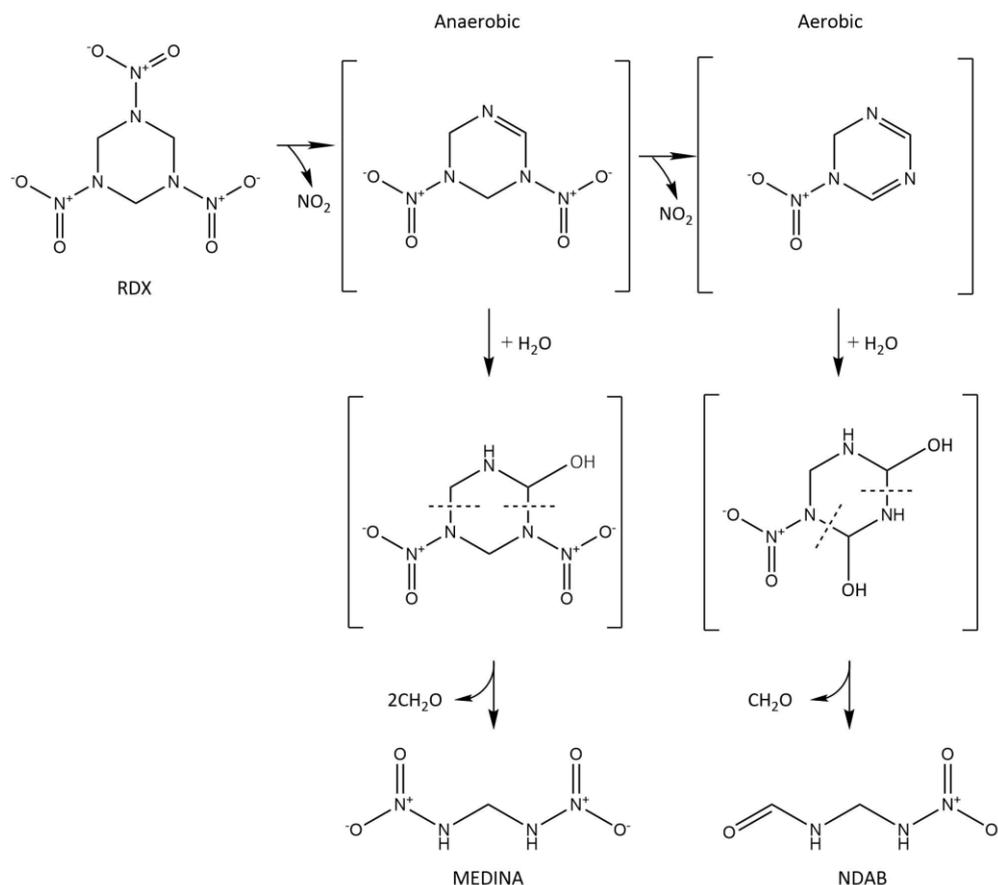


**Figure 5.1:** Scheme illustrating the bio-conversion of xenobiotic compounds in plant cells. In Phase I, the xenobiotic (orange hexagon, X) in the cytosol is transformed into a generally more soluble derivative via hydrolysis, oxidation or reduction (blue circle). In phase II, the derivatised compound is conjugated with glutathione (yellow trapeze, G), with malonate (violet triangle, M) or with a molecule of sugar (gray hexagon, S) by GST (glutathione S-transferase), UGT (glucosyl transferase) or Mt (malonyl transferase), respectively. In phase III the conjugated molecules are transported into the vacuole or in the cell wall.

Phytoremediation exploits the natural capability of plants to take up compounds that are present in their surroundings and then detoxify them. This ability, coupled by the presence of endogenous metabolic systems, can be further amplified by the introduction, via genetic engineering, of external genes, with known activity towards the target compounds, leading to a more efficient biotransformation and possibly less toxic derivatives. Successful examples of engineered P450-based systems for phytoremediation purposes are numerous. Kawahigashi *et al.* developed transgenic rice (*O. sativa*) crops (named pKABCH, from rice cv. Nipponbare) with increased ability to extract numerous herbicides, from different media, due to the tandem expression of human CYP1A1, CYP2B6, and CYP2C19 (59). Mammalian CYP2E1, due to its demonstrated activity towards a broad array of environmental contaminants, such as trichloroethylene (TCE), benzene and chloroform (231), has attracted major attention over the years, leading to the development of multiple transgenic systems. Starting from assays with modified tobacco plants (232), Doty *et al.* created a CYP2E1-expressing poplar hybrid (*Populus tremula* x *P. alba*, clone INRA 717-1B4), which was able to extract benzene and TCE from the air, and chloroform from a hydroponic solution (233). Subsequently, alfalfa plants (*Medicago sativa* L.) were modified by Zhang and coworkers to express human CYP2E1 and a GST. The resulting plants displayed enhanced cross-resistance to mixtures of xenobiotics (such as TCE, cadmium and mercury, if compared to wild type plants (234, 235).

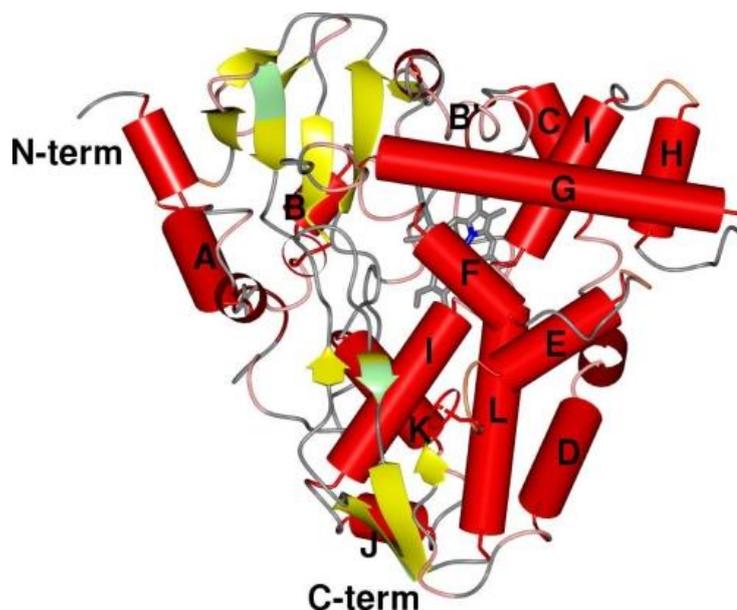
Besides volatile organic compounds and toxic metals, phytoremediation processes have been developed also for the removal and detoxification of explosives pollution, and, given the widespread nature of the contamination, a more effective alternative to the conventionally-used incineration, soil capping, soil removal or composting techniques (236). Prof. Neil Bruce's group particularly focused extensive research efforts towards the development of viable phytoremediation strategies for the removal of nitroexplosives, such as the widely-used 2,4,6-trinitrotoluene (TNT) and hexahydro-1,3,5-trinitro-1,3,5-triazin (RDX). Both compounds, characterised by a ring with multiple nitro groups (figure 5.2), are toxic and recalcitrant to degradation in the environment.





**Figure 5.4:** Illustration of the XplA (CYP177A1) mechanism of action towards RDX, as proposed by Jackson *et al.* (237), on the basis of the measured nitrite and formaldehyde released in anaerobic and aerobic conditions. In anaerobic conditions, one mole of nitrite and two moles of formaldehyde are formed; under aerobic conditions, two moles of nitrite and one mole of formaldehyde are formed. In addition, in absence of oxygen MEDINA (methylenedinitramine) is formed, whereas when oxygen is present, NDAB (4-nitro-2,4, diazabutanal) is formed. The intermediates in brackets are hypothetical.

Homologues of XplA (with more than 90 % amino acid sequence identity on a region covering > 60 % of the enzyme) were also found to be expressed by other bacterial isolates from RDX-contaminated soils in Australia, Israel and Belgium (88). Further activity and structural studies (figure 5.5) were conducted by Dr. Sabbadin. Assays performed with the purified XplA heme showed a  $K_m$  for RDX of  $10.23 \pm 0.90 \mu\text{M}$  (91, 238).

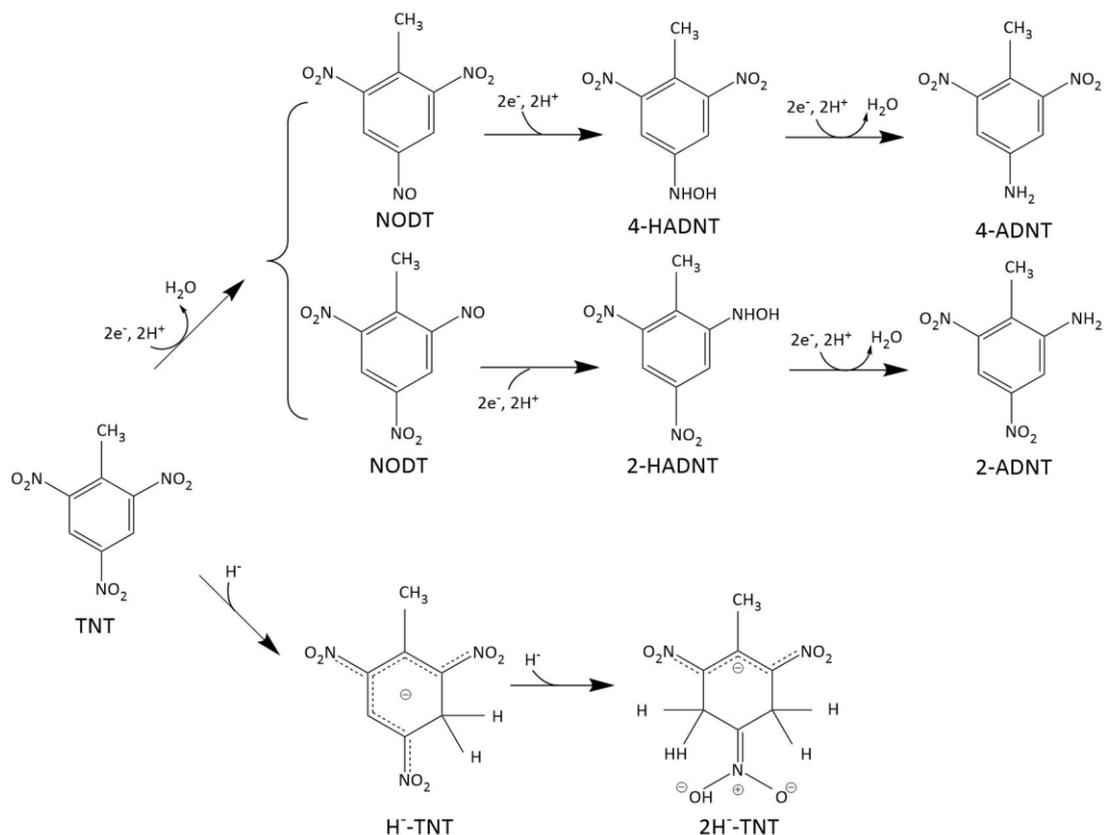


**Figure 5.5:** Schematic representation of the elucidated structure of the XplA heme domain (figure from Sabbadin *et al.* (238)).

The applicability of this enzymatic system for the phytoremediation of RDX was tested by introducing XplA into *Arabidopsis* (*A. thaliana*). *Arabidopsis* lines expressing the XplA heme were able to remove RDX from saturating solutions. Subsequent soil studies demonstrated that the transgenic lines were more resistant to RDX and were able to use it as a source of nitrogen to enhance biomass production. (88). These properties were further enhanced when XplB was also transformed into the plant (237).

The explosive TNT is more recalcitrant to degradation than RDX. Studies over the years reported that biological activity towards TNT consists mainly of reductive reactions. One pathway sees the sequential conversion of TNT to the highly unstable 4-nitroso-2,6-dinitrotoluene (NODT) which is reduced to 4-hydroxyl-2,6-dinitrotoluene (HADNT), and then aminodinitrotoluenes (ADNT), by nitroreductases in bacteria (239). In plants, OPRs (163) are known to contribute to these reductive transformation steps, then both HADNTs and ADNTs can be conjugated, by UGTs to glucose molecules. Denitration with direct conjugation of GSH to the ring can also occur, catalysed by GSTs (164). The TNT conjugates are then thought to be sequestered into the cell wall or in the vacuoles (162). Another pathway consists of

the reduction of the TNT ring, yielding the hydride/dihydride-Meisenheimer complexes (figure 5.6). In bacteria, these complexes can condense, forming diarylamines (240).



**Figure 5.6:** Steps in the biotransformation of TNT in plant roots and in bacteria, as described by Rylott *et al.* (241). Multiple consecutive reduction steps of the nitro groups can lead first to the formation of the unstable 2/4-nitroso-4/2,6-dinitrotoluene, then to the 2/4-hydroxylamino-derivative and finally to 2/4-amino-4/2,6-dinitrotoluene. The latter two types of derivatives can be further glycosylated by UGT transferases. Alternatively, through reduction of the ring, TNT is converted into the hydride/dihydride-Meisenheimer complexes. In bacteria, the condensation of these complexes with HADNT can occur, with formation of diarylamines. In addition, GST transferases can conjugate the TNT molecule to glutathione. NODT = 2/4-nitroso-4/2,6-dinitrotoluene; HADNT = 2/4-hydroxylamino-4,6-dinitrotoluene; ADNT = 2/4-amino-4/6-dinitrotoluene;  $H^-/2H^-$  TNT = hydride/dihydride-Meisenheimer complex.

At the end of all these reductive transformations steps, affecting specifically the functional groups of the molecule (nitro,  $-NO_2$ ) the aromatic ring remains intact. Besides some cases seen in fungi, studies on the mineralisation of TNT are quite limited, due to the high stability of the aromatic ring, whereby the carbon bonds are sterically 'protected' from attack (242).

Mineralisation levels observed in multiple bacterial strains exposed to TNT were minimal (< 1 %, if any) (243, 244). In the case of fungi, for example *P. chrysosporium*, *Irpex lacteus* and *Clitocybula duseinii* b12, TNT was firstly reduced to the amino derivatives and further metabolised, with release of water and CO<sub>2</sub> (245-247). This kind of activity, absent in bacteria, was assigned to fungal endogenous enzymes, involved in the ligninolytic pathway, such as laccases, manganese peroxidases and lignin peroxidases (248, 249).

In contrast to the previously described reductive mechanisms (seen in figure 5.6), Bhadra *et al.*, investigated oxidative TNT metabolic pathways. These studies were conducted with the aquatic plant parrot feather (*Myriophyllum aquaticum*), incubated in TNT-containing liquid medium. The HPLC-MS coupled with NMR analyses revealed the formation of a set of metabolites, all still presenting an intact aromatic ring, such as 2-amino-4,6-dinitrobenzoic acid (2-ADNBA), 2,4-dinitro-6-hydroxy-benzyl alcohol, acetoxymino-4,6-dinitrobenzaldehyde and 2,4-dinitro-6-hydroxytoluene (165). The metabolite 2-ADNBA alongside 2,4,6-trinitrobenzoic acid (TNBA) were also observed in soil samples taken from a TNT-contaminated ammunition plant in Germany (250).

A number of cytochrome P450 enzymes were among the hundreds of upregulated genes revealed in multiple gene expression studies performed on Arabidopsis plants exposed to TNT (160-162). Among the overexpressed genes were also TNT-active GSTs, UGTs and OPRs. The most upregulated cytochrome P450 was CYP81D11, with expression levels exceeding the basal expression (without TNT) by more than 25 fold, and at comparable expression levels with many of the TNT-active genes. Furthermore, gene co-expression data regarding CYP81D11 on the CYPedia platform (251) demonstrated that, in multiple cases, CYP81D11 is expressed in the same tissues or overexpressed upon the same stimuli as glucosyl transferases, toxin/xenobiotic metabolisers and stress response agents (all co-expression data sets for CYP81D11 can be visualised at: [http://www-ibmp.u-strasbg.fr/~CYPedia/CYP81D11/CoExp\\_CYP81D11\\_Pathways.html](http://www-ibmp.u-strasbg.fr/~CYPedia/CYP81D11/CoExp_CYP81D11_Pathways.html)).

The role of CYPs in the detoxification of TNT has not been investigated to date. Given the multiple evidences described above, including the discovery of oxidative derivatives of TNT in plant incubations, the overexpression of CYP enzymes upon treatment with TNT, as well as the gene co-expression data, it is hypothesised that the upregulated CYPs have a role in the detoxification of TNT.

The aim of the work described in this chapter is to further our understanding of the function of CYP81D11, the most upregulated CYP upon TNT treatment, and to unravel new potential ways to detoxify this toxic environmental pollutant.

## 5.2 Materials and Methods

### 5.2.1. Plants and growth media

#### Plants

Tobacco (*Nicotiana benthamiana*) seeds were kindly donated by Luisa Elias, CNAP, University of York.

*Arabidopsis* (*Arabidopsis thaliana*) *CYP81D11*-knockdown lines and *CYP81D11*-overexpressing lines were kindly provided by Prof. J. Napier, Rothamsted Research. The wild type line, used as a control, was ecotype Columbia Col0.

#### Growth media

Soil: Levingtons F2 compost, treated with the Intercept insecticide (Imidacloprid), according to the manufacturer's instructions, as a prophylaxis against sciarid fly larvae.

½ MS (Murashige and Skoog) medium (132): 2.15 g/l of MS medium, 20 mM sucrose, dissolved in distilled ELGA-pure water. Plant agar (1% w/v) was added for the preparation of ½ MS agar plates. pH was adjusted to 5.7 with NaOH prior to autoclave sterilisation.

Tobacco plants were grown in soil for up to six weeks, prior to leaf disc harvest.

Arabidopsis plants used for the incubation assays were grown on sterile ½ MS agar and subsequently transferred in ½ MS liquid medium.

### **5.2.2. Sterilisation of Arabidopsis seeds**

The seeds were sterilised with the vapour-phase method. Briefly, small aliquots of seeds (~70) in open 1.5 ml Eppendorf tubes were exposed to the fumes from 100 ml of bleach mixed with 3 ml of hydrochloric acid, in an airtight container placed in a fume hood. After 4 h, the container was vented within the fume hood, and left for further 10 min to release the residual toxic gas.

### **5.2.3 Verification of gene expression in the Arabidopsis *CYP81D11*-modified lines via qPCR**

The RNA extraction was performed following the procedure used by Dr. Lorenz (161) and using an ISOLATE II RNA Plant Kit (Bioline). Plant tissues (harvested from whole, liquid-culture-grown two-week old plants, incubated for 6 h with 60 µM TNT) were frozen in liquid nitrogen, then ground using a pestle and a mortar. The RLY lysis buffer, from the kit, mixed with 10 % (v/v) β-mercaptoethanol, was added to the tissue powder (350 µl/100 mg of tissue). After a brief vortex, the lysate was filtered through a spin column (provided in the kit). The flow-through was then mixed 1:1 (v/v) with 70% (v/v) ethanol and loaded onto the RNA plant column. After a washing and a desalting step, the sample retained on the silica membrane was treated for 15 min with DNaseI. The membrane was then subjected to three washing steps and the RNA eluted via centrifugation with RNase-free water. The purity of the resulting RNA was checked using a nanodrop spectrophotometer, as good quality samples have an Abs<sub>260</sub>/Abs<sub>280</sub> ratio of approximately 2. The reverse transcription from the extracted RNA to cDNA was carried out employing a SuperScript II reverse transcriptase kit (Invitrogen). Following the instructions from the manufacturer, the reaction mixture consisting of 1 ng - 5 µg of RNA, oligo(dT)<sub>12-18</sub> and dNTPs was firstly incubated at 65 °C for 5 min. The buffer (provided in the kit) and 0.1 M DTT were subsequently added

and the mix was cooled to 42 °C for 2 min. After the addition of the reverse transcriptase, the reaction was carried out at 42 °C for 50 min. At the end of the incubation, the enzyme was inactivated at 70 °C for 15 min. The resulting single-strand cDNA was used as a template for the quantitative PCR amplification executed on a StepOne Plus Real Time PCR Systems machine (Applied Biosystems), using specifically designed oligonucleotide primers. Oligos complementary with the *ACTIN2* gene (At3g18780) were used for control amplification reactions (all primers are listed in table 5.1), to normalise the data collected. The measurements were performed with five biological replicates, each in technical triplicate.

**Table 5.1:** List of oligonucleotides used for the measurement of the gene expression in the Arabidopsis modified plants. Actin was used as an endogenous control.

Target gene	Oligo sequence
<b><i>CYP81D11</i></b>	Forward 5' – TTATGATACTTGCCGGGACTG – 3'
	Reverse 5' – TCGATTTCCGGTCTTTGCC – 3'
<b>Actin</b>	Forward 5' – TACAGTGTCTGGATCGGTGGTT – 3'
	Reverse 5' – CGGCCTTGAGATCCACAT – 3'

The use of SYBR® fluorescent dye (ThermoScientific) allowed the detection of the synthesis of double-stranded DNA products. The resulting melt curve data were collected and analysed with the StepOne™ Software.

#### **5.2.4. Hydroponic cultures of Arabidopsis transgenic plants & HPLC analysis of TNT uptake**

Surface-sterilised seeds were placed on ½ MS agar plates (in single rows of 20 seeds for the root-length experiments), subsequently wrapped in aluminum foil. The plated seeds were stratified at 4 °C for 72 h, then placed in a growth chamber with a 16-h photoperiod, 21 °C day and 18 °C night temperatures and 100 µmol/m<sup>2</sup> s of light.

After one week, germinated seedlings, in batches of 8 seedlings, were transferred from the plates into 100 mL Erlenmeyer flasks containing 20 ml  $\frac{1}{2}$  MS medium, supplemented with 20 mM sucrose, pH 5.7. The flasks were placed on a rotating shaker at 125 rpm, under 20  $\mu\text{mol}/\text{m}^2 \text{ s}$  of light. After two weeks of growth, the medium was replaced with an aqueous solution of 200  $\mu\text{M}$  TNT. Samples from the liquid medium were collected at regular time points for HPLC analysis. After a centrifugation step for 15 min at 34,000 g, the samples were subjected to HPLC analysis (Waters X-Bridge C18 column, column temperature: 30 °C, mobile phase: 48 % methanol / 52 %water).

### **5.2.5. Root length on-plate comparison for differential TNT-resistance**

Germinated seeds were transferred onto  $\frac{1}{2}$  MS plates (20 seedlings per plate), supplemented with 20 mM sucrose and different concentrations of TNT, ranging from 0 to 30  $\mu\text{M}$  and grown vertically in the same chamber described in section 5.2.4. Pictures of the plates were taken after 20 days and the length of the roots were measured with the ImageJ processing program (252).

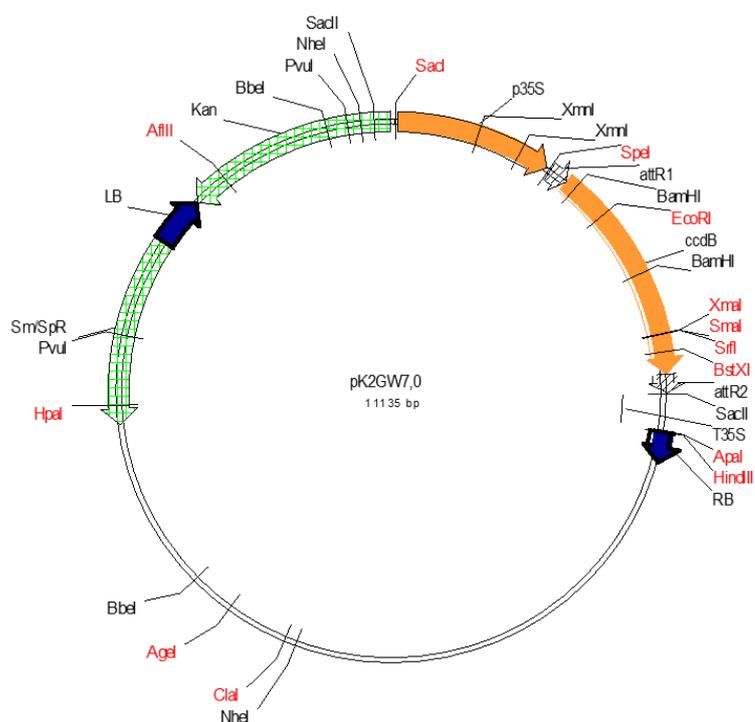
### **5.2.6. Transient expression of CYP81D11 in *Nicotiana benthamiana***

#### **Cloning**

The cDNA extracted from Arabidopsis whole seedlings was used as template for the amplification of the genes of interest, employing gene-specific primers. A second PCR amplification was performed in order to add the adapter ends for the “LR” recombination step of the Gateway cloning process (253). The two sets of primers used for the sequential PCR reactions are listed in table 5.2. The original Gateway cloning protocol was slightly modified, as 100 ng of purified PCR amplicons were added to 150 ng of the pK2GW7 destination vector (figure 5.7) in TE buffer, pH 8.0.

**Table 5.2:** Oligonucleotides used for the 3-step amplification of the target gene for cloning in the pK2GW7 plasmid. PCRExt: 5'-3' gene-specific primers, with start codon and stop codon in bold / PCRHis: primers for the introduction of the His tag / PCRAttL: primers for the introduction of the LR recombination sites.

Name	Oligo sequence
PCRExt	Forward 5' - CCAACTTTGTACAAAAAGCAGGCT <b>ATG</b> TCATCAACAAAGACAATAATG – 3'
	Reverse 5' - <b>TTA</b> GTG GTG GTG GTG GTG GTG TGG ACA AGA AGC ATC TAA AAC – 3'
PCRHis	Forward 5' - CCAACTTTGTACAAAAAGCAGGCT <b>ATG</b> TCATCAACAAAGACAATAATG – 3'
	Reverse 5' - CCA ACT TTG TAC AAG AAA GCT GGG <b>TTT</b> AGT GGT GGT GGT GGT
PCRAttL	Forward 5' - CAAATAATGATTTTATTTTGACTGATAGTGACCTGTTTCGTTGCAACAAA TTGATAAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCA – 3'
	Reverse 5' - CAAATAATGATTTTATTTTGACTGATAGTGACCTGTTTCGTTGCAACAAA TTGATAAGCAATGCTTTCTTATAATGCCAACTTTGTACAAGAAAGCT – 3'



**Figure 5.7:** Map of the pK2GW7 vector employed for the expression of CYP81D11 in the tobacco plants.

The reaction mixture, after the addition of 2  $\mu$ l of LR Clonase II enzyme mix (ThermoFisher Scientific), was incubated for 1 h at 25  $^{\circ}$ C. The reaction was terminated by adding 1  $\mu$ l of proteinase K solution and further incubated at 37  $^{\circ}$ C for 10 min.

Competent DH5 $\alpha$  *E. coli* cells were transformed with the reaction mixture. Colonies were picked and inoculated in LB medium for overnight incubation at 37 °C and subsequent plasmid isolation and sequencing.

#### **Agrobacteria electroporation and cultures**

The purified plasmids carrying the *CYP81D11* gene, along with the empty pK2GW7 vector and the P19 construct (viral suppressor of gene silencing) (254), were transformed via electroporation into *A. tumefaciens* competent cells (for complete protocol see Materials and Methods, paragraph 2.4.7.2). After 48-72 h of incubation at 30 °C on LB agar plates (supplemented with 50  $\mu$ g/ml spectinomycin and 50  $\mu$ g/ml gentamycin for selection), single colonies were picked and inoculated into 100 mL of LB medium with antibiotics and incubated again at 30 °C with shaking.

#### **Infiltration and harvesting of leaf discs**

The Agrobacterium cultures were harvested 48 h later via centrifugation, at 5,000 g for 15 min and washed twice with a ½ volume of double distilled water. The density of the cultures was checked spectrophotometrically and adjusted to a final OD<sub>600</sub> value of 0.4. CYP cultures were mixed 1:1 with P19 cultures for co-expression and infiltrated using a 1-ml syringe into the abaxial side of the leaf of 5-6 week-old tobacco plants. Cell cultures harboring the empty vector were also mixed with P19 cultures and infiltrated into tobacco leaves, as negative controls. Three leaves per plant were infiltrated, in a total of 8 plants for the CYP81D11 cultures and 8 for the empty vector cultures. After 5 days, 20 discs (approximately 0.5 cm diameter) from each plant were excised from the treated leaves. Half of the harvested discs (10) were subjected to treatment with TNT whereas the remaining 10 were collected for subsequent verification of protein expression through SDS-PAGE and western blot analysis.

### **Verification of protein expression: protein extraction, SDS-PAGE verification and western blot analysis**

Approximately 53 mg of tissue were obtained from each infiltrated leaf and ground in liquid nitrogen. The resulting powder was resuspended in Lysis buffer (50 mM Tris-HCl pH 7.8, 10 mM EDTA, 300 mM sucrose, 100 mM NaCl, 1 mM DTT and 200  $\mu$ M PMSF). The samples were then subjected to sequential centrifugations, 10 min at 1,000 g, 7 min at 2,000 g and 5 min at 1,000 g to remove tissue residues. For the complete SDS-PAGE and western blot procedures see Materials and Methods, paragraphs 2.5.4 and 2.5.5.

### **Leaf discs incubations**

Ten leaf discs from the treated plants (CYP81D11 cultures or empty vector cultures), along with discs from non-infiltrated plants (second negative control set) were incubated in 20 mL of  $\frac{1}{2}$  MS medium containing 100  $\mu$ M TNT. The discs were soaked by vacuum application. The 100 ml glass Erlenmeyer flask containers were wrapped in aluminum foil, to minimise the photodegradation of TNT. Samples of the liquid medium as well as of the tissues were collected over a period of 48 h. In order to detect the levels of TNT uptake and TNT-derivatives in the CYP-expressing samples and the controls, the medium samples were centrifuged for 10 min at 13,000 rpm and the supernatants analysed by HPLC (Waters X-Bridge C18 column, column temperature: 30 °C, isocratic mobile phase: 50% water / 50% acetonitrile). With this method, TNT eluted at  $\sim$  8.2 min and peak areas were integrated at 230 nm.

### **Metabolite extraction from the leaf discs**

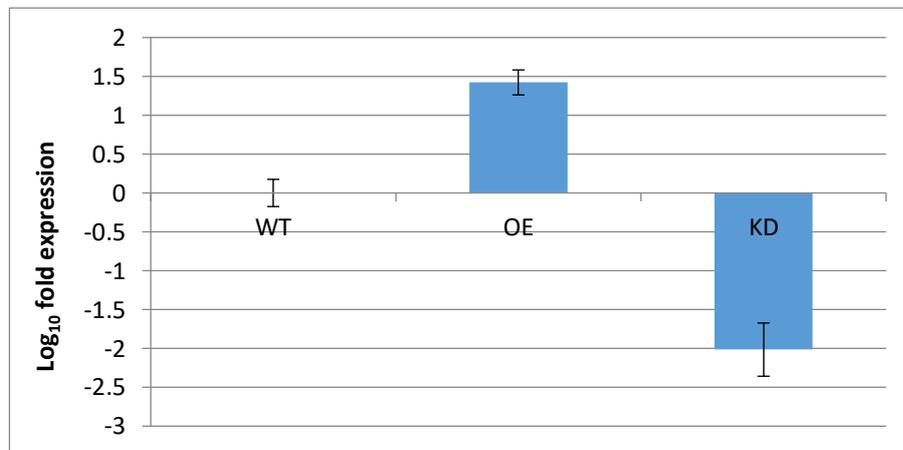
Following the procedure used by Hannink *et al.* (239), the TNT metabolites retained in the tissues were extracted by grinding the discs in liquid nitrogen. The resulting powder was re-suspended in acetonitrile (10 ml / g of tissue). After two vortex spins (10 s each), the tissue debris was pelleted via centrifugation at 14,000 rpm and the acetonitrile evaporated with a GeneVac E22 evaporator (SP scientific), for 45 min at 60 °C. Subsequently, the pellets were re-suspended in 400  $\mu$ l acetonitrile and analysed via HPLC (Waters X-Bridge C18 column, column temperature: 30 °C, mobile

phase: 50% acetonitrile / 50 % 20 mM tetrabutylammonium dihydrogen phosphate (TBAP), pH 7).

## 5.3 Results

### 5.3.1 Verification of CYP81D11 expression in Arabidopsis

The modified levels of CYP81D11 transcripts in the Arabidopsis lines received from Rothamsted Research were quantified via qPCR expression analysis. The results of the quantitative  $\Delta\Delta C_t$  comparative analysis, reported in figure 5.8, confirmed that upon treatment with TNT the CYP81D11-modified plants produced higher levels (for the *CYP81D11*-overexpressors) or lower levels (*CYP81D11*-knockdown) of the transcript, when compared to the wild type plants.

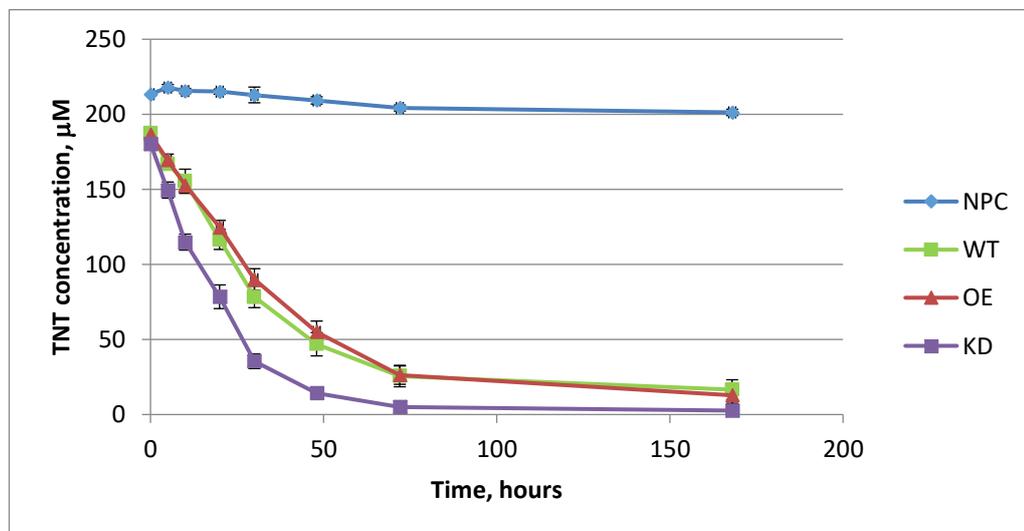


**Figure 5.8:** qPCR results from the cDNA obtained from the two-week-old Arabidopsis plants grown on ½ MS medium supplemented with 60  $\mu$ M TNT. WT = wild type, OE = *CYP81D11*-overexpressor, KD = *CYP81D11*- knockdown. The results shown are the average of 5 biological replicates  $\pm$  S. E.

The qPCR data collected for the untreated plants were not clear, as no significant difference in expression levels could be observed between the wild type and *CYP81D11*-knockdown plants (Log<sub>10</sub>(fold expression) values: wild type =  $-1.89 \pm 0.12$  / *CYP81D11*-overexpressor =  $3.27 \pm 0.13$  / knockdown =  $0.56 \pm 0.24$ ).

### 5.3.2 TNT uptake by hydroponic Arabidopsis cultures

The result of the HPLC analysis of the samples collected from the plant growth medium supplemented with TNT is shown in figure 5.9.

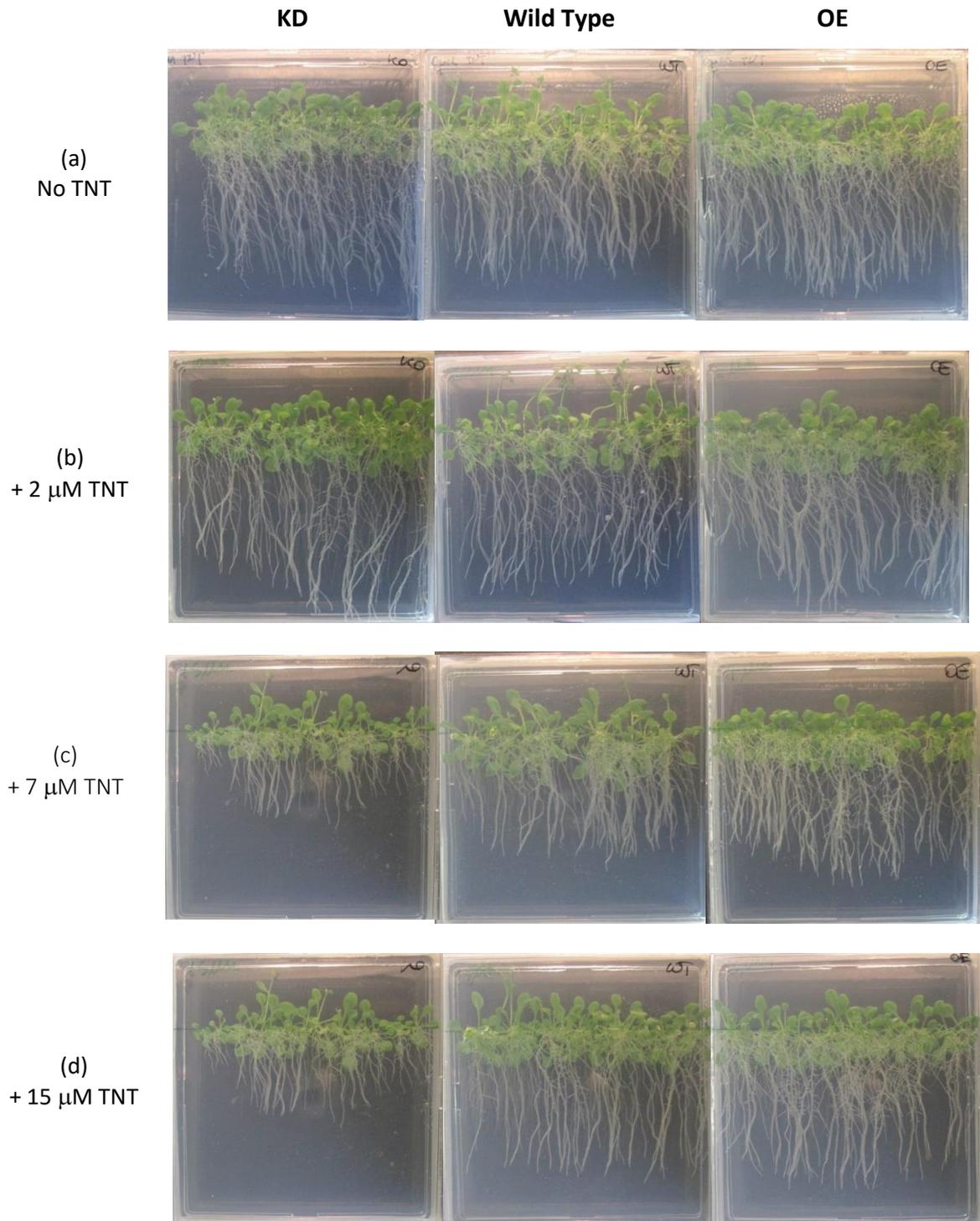


**Figure 5.9:** TNT concentration in the plants' growth medium. The 3-week-old seedlings were incubated with 200 µM TNT in ½ MS for 1 week. NPC = no plant control, WT = wild type control, OE = *CYP81D11*-overexpressor, KD = *CYP81D11*-knockdown. Results shown are means of 8 biological replicas ± SE. Statistical analysis carried out with SPSS OneWay Anova,  $p < 0.05$ .

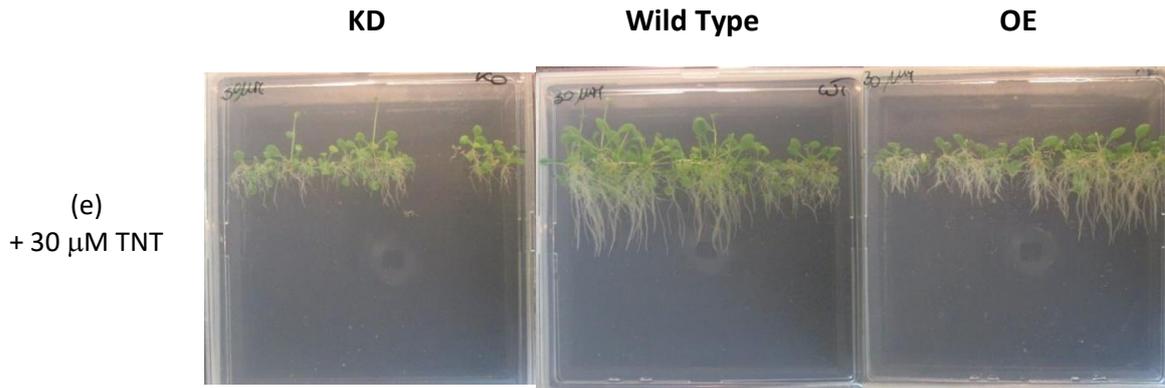
The HPLC analysis revealed differences in TNT uptake of the liquid-culture grown wild type and *CYP81D11*-modified plants. The knockdown plants showed a significantly faster uptake (for all the recorded values up to the 48 h time point, verified with SPSS OneWay Anova,  $P < 0.05$ ) of the xenobiotic, than the wild type or overexpressing line. For the knockdown line, more than 90 % of the TNT was removed from the medium in the first 48 h; whereas, for the overexpressing and wild type lines, 7 days were needed to remove the same amount of TNT. As expected, in the no plant control samples the concentration of TNT remained stable throughout the time frame.

### 5.3.3 Comparison of *CYP81D11*-modified plants' root length

To test for a possible correlation between the expression of *CYP81D11* and resistance to TNT, a comparative assay was carried out, by growing the wild type, *81D11*-knockdown and *81D11*-overexpressing Arabidopsis plant lines on TNT-containing plates. Pictures of the plates were taken after 20 days (figure 5.10 a-e).

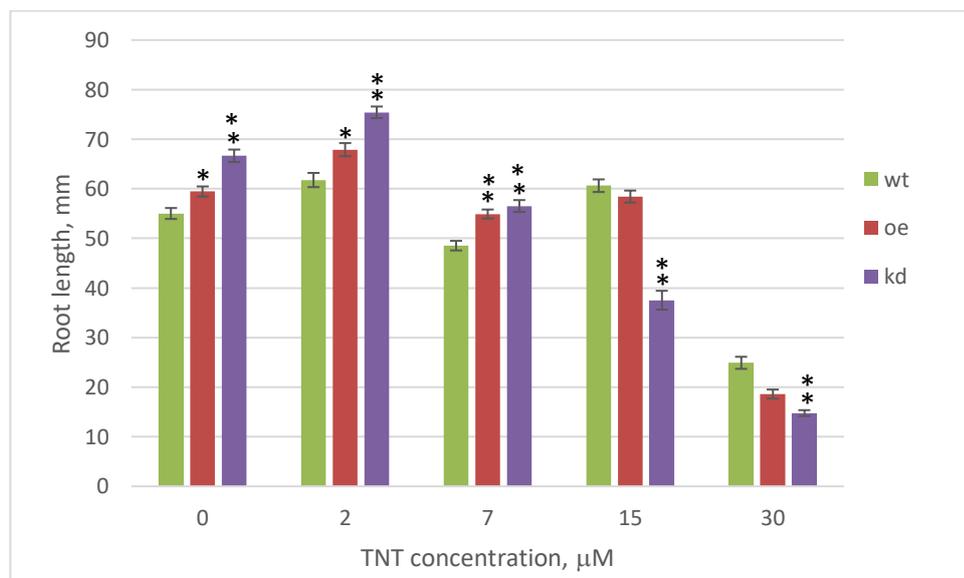


(Figure 5.10 a-e, continues on following page)



**Figure 5.10 a-e:** Pictures of 20-day-old Arabidopsis plants grown on 1/2MS + sucrose plates. Each row (a-e) represents one TNT concentration set, in the range 0 - 30  $\mu$ M. Left column: KD (*CYP81D11*-knockdown) / middle column: Wild type / right column: OE (*CYP81D11*-overexpressor).

The root measurements are shown in figure 5.11. A specific trend could be observed across the lower concentrations of TNT (0, 2, 7  $\mu$ M): both overexpressing and knockdown lines displayed longer roots, compared to wild type. When a higher concentration of TNT was present (15, 30  $\mu$ M) in the plate, the trend was reversed, with the knockdown plants displaying significantly shorter roots, in comparison to the wild type and overexpressing lines.

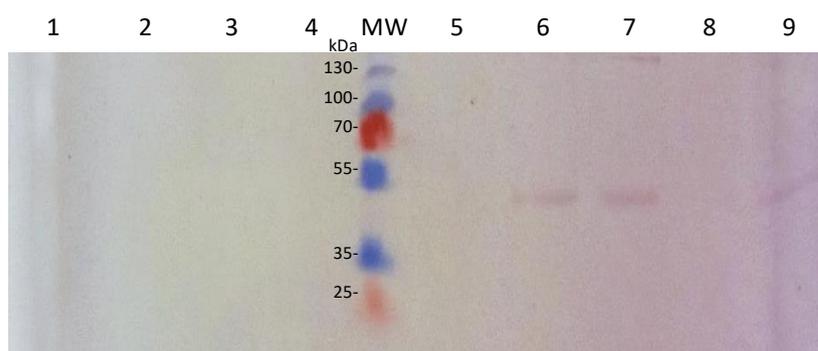


**Figure 5.11:** Root lengths of 20-day old Arabidopsis seedlings grown on plates with different concentrations of TNT. Three replicates for each concentration and line, 20 seeds per plate. WT = wild type, OE = *CYP81D11*-overexpressor, KD = *CYP81D11*-knockdown. Statistical analysis performed with OneWay Anova and Tukey test. One asterisk (\*) denotes data sets significantly different from the wild type, with  $p < 0.05$ ; two asterisks (\*\*) denote data sets significantly different from WT, with  $p < 0.01$ .

### 5.3.4 Transient expression of CYP81D11 in tobacco leaves

#### Expression of *CYP81D11*

Following the established protocol used by colleagues at the Institut de Biologie Moléculaire des Plantes (IBMP, CNRS), the construct containing the histidine-tagged *CYP81D11* native sequence was transformed into *Agrobacterium*. Transformed cultures, mixed with cultures expressing the viral P19 inhibitor, were infiltrated into tobacco leaves. Half of the discs excised from the leaves were used for the verification of protein expression (figure 5.12) and the other half for TNT incubations.

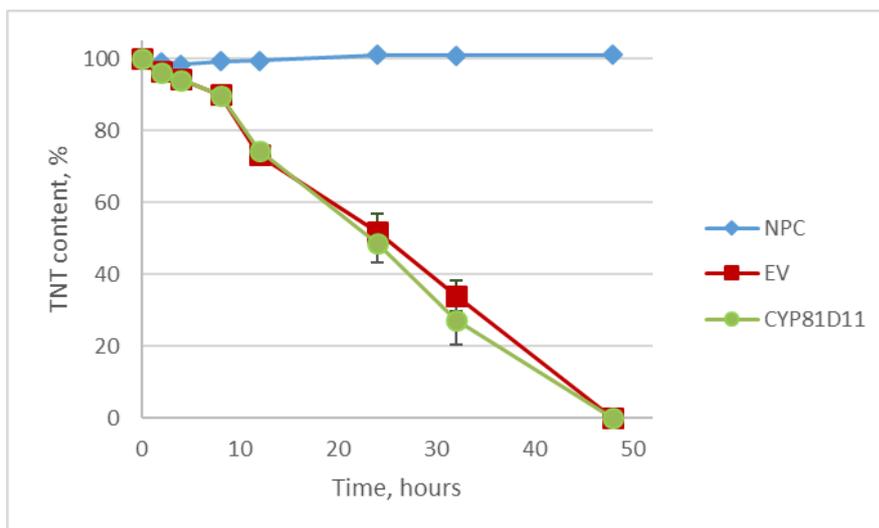


**Figure 5.12:** Western blot analysis of CYP81D11 expression in the infiltrated tobacco leaves (samples 1-4: empty vector control / samples 5-9: CYP81D11 samples).

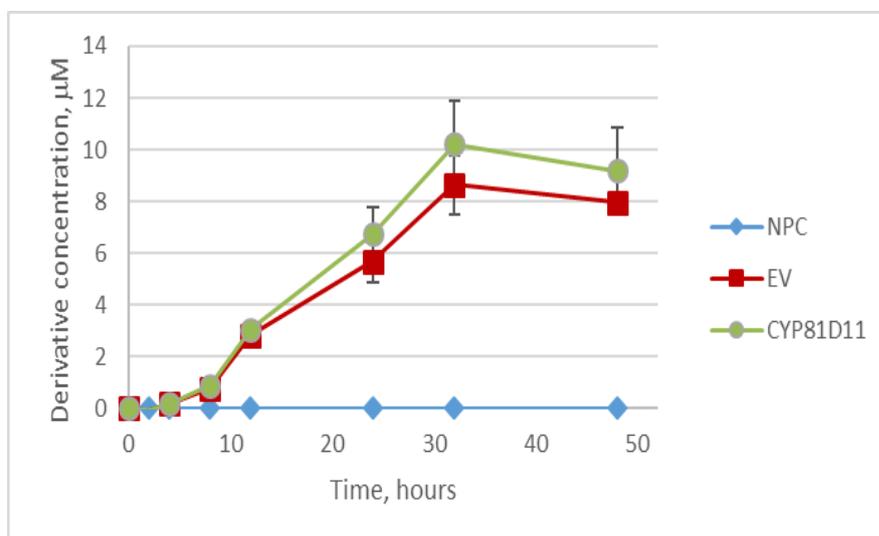
Three of the samples examined via western blot clearly gave a band at the expected size (lanes 6, 7 and 9); for the other two samples (lanes 5 and 8), the concentration of the his-tagged protein might have been too low for detection.

#### Analysis of TNT uptake and conversion

Time point samples from the tobacco leaf discs incubations dosed with TNT were analysed via HPLC. All the resulting chromatograms (except the no plant controls) presented two peaks, one eluting at 6.2 min and the other at approximately 8.3 min. Analysis of the standards demonstrated that the 6.2' peak corresponded to a mixture of 2/4-ADNTs, while the 8.3' peak corresponded to the TNT standard. The collected data are shown in figures 5.13 and 5.14.



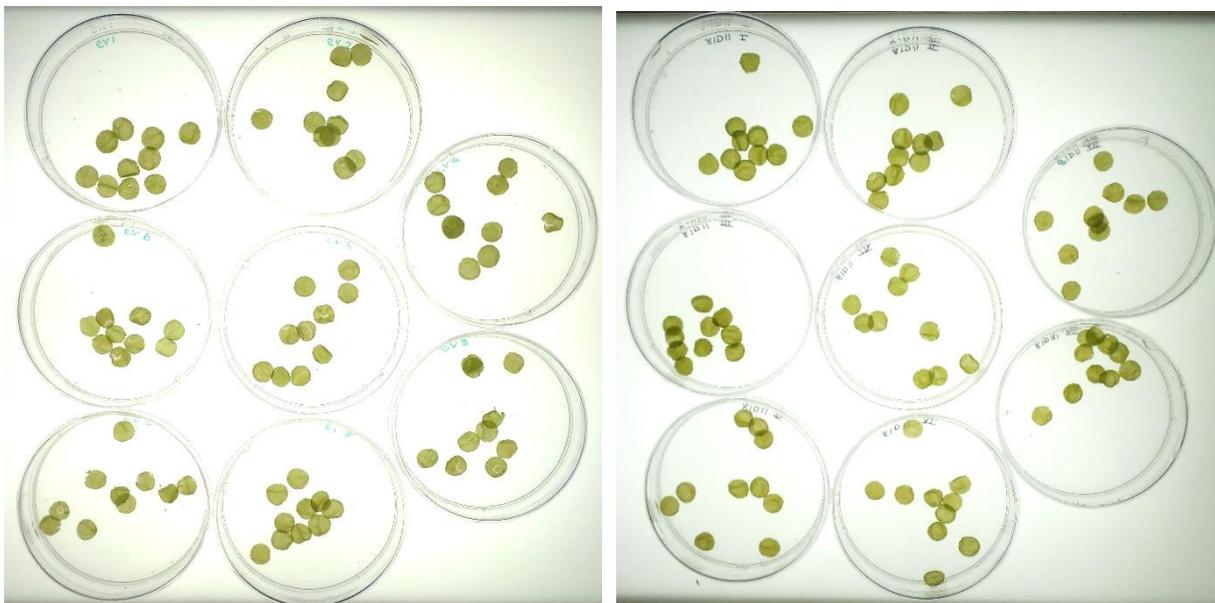
**Figure 5.13:** TNT uptake observed during the incubation of tobacco leaf discs expressing CYP81D11 on ½ MS + 20 mM sucrose containing 100 µM TNT. The samples were analysed at the HPLC (Waters C18 XBridge column, 5 µm, 4.6 x 250 mm) using an isocratic method, with the mobile phase consisting of 50:50 water:acetonitrile. NPC = no plant control, EV = empty vector. Mean of eight biological replicates ± s.e.



**Figure 5.14:** Levels of ADNT derivatives observed during the incubation of tobacco leaf discs expressing CYP81D11 on ½ MS + 20 mM sucrose containing 100 µM TNT. The samples were analysed at the HPLC (same method as above) NPC = no plant control, EV = empty vector. Mean of eight biological replicates ± s.e.

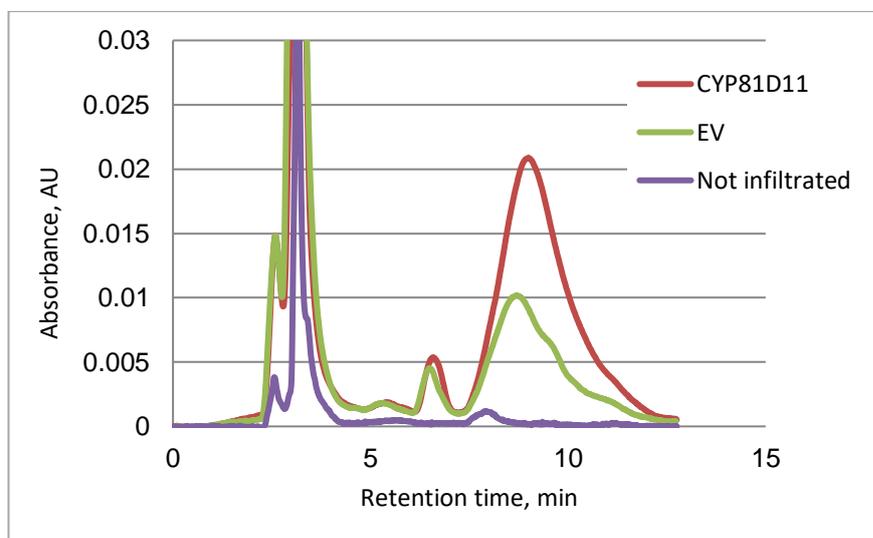
The analysis of the samples of the liquid medium showed that there was no significant difference (verified with a statistical T-test) in the rate of TNT removal nor in the formation of TNT derivatives by the discs expressing CYP81D11, when compared to those transformed with the empty vector alone.

At the end of the incubation period, the leaf discs were collected. Upon visual inspection no significant difference could be noticed, in terms of chlorosis (figure 5.15).



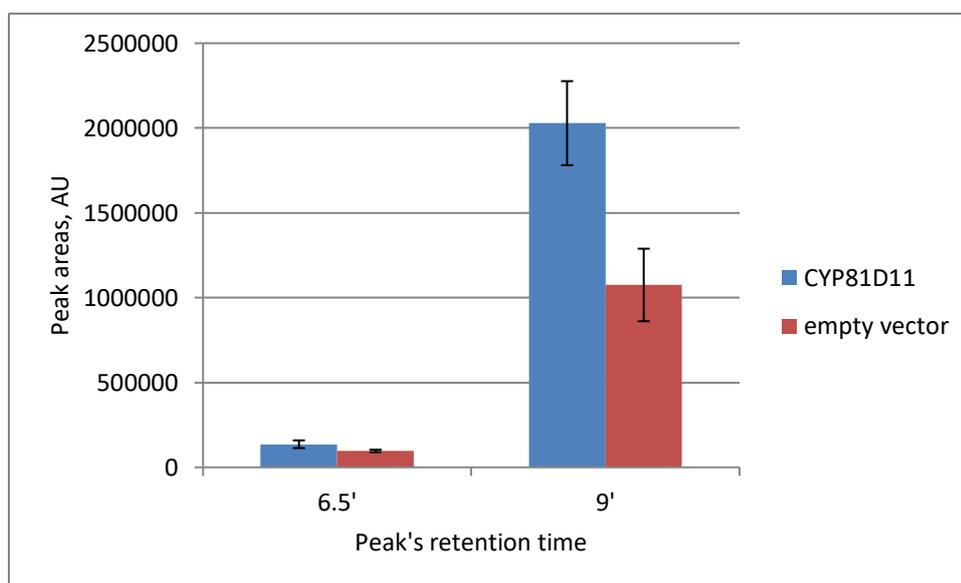
**Figure 5.15:** Pictures of the tobacco leaf discs after 48 h of incubation on  $\frac{1}{2}$  MS + 20 mM sucrose, in presence of 100  $\mu$ M TNT. Left panel: empty vector control/ right panel: CYP81D11-expressor. Each dish contains the ten discs harvested from each biological replicate.

Levels of TNT metabolites found in the tissues of the wild type, knockdown and overexpressing lines are shown in figures 5.16 and 5.17.



**Figure 5.16:** Chromatographic profiles of the samples collected from the tobacco leaf tissues. The metabolites were extracted with acetonitrile and analysed via HPLC, employing an isocratic method 50:50 20 mM TBAP pH 7:acetonitrile. EV = empty vector control.

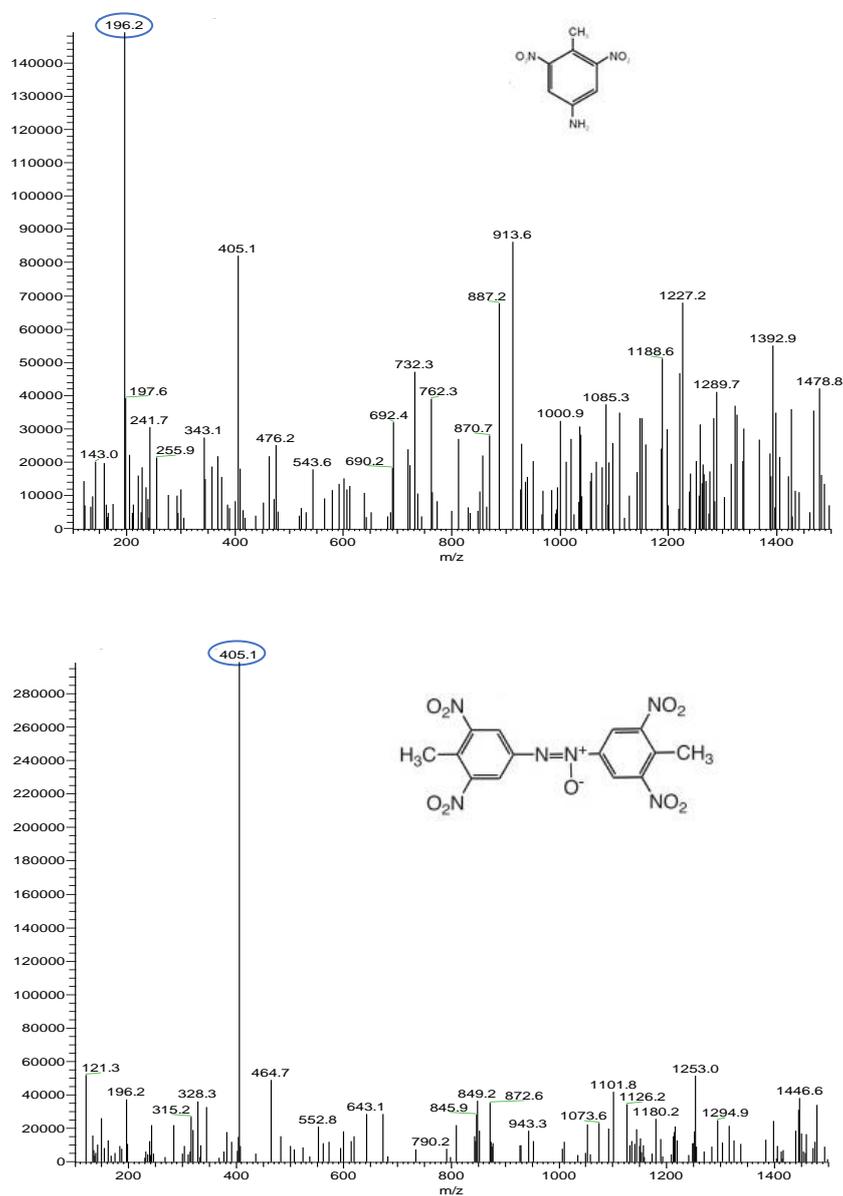
The TNT standards (25, 50, 100  $\mu\text{M}$ ) were subjected to the same treatment as the ground tissues (mixed with acetonitrile, GeneVac evaporation and resuspension), in order to verify whether TNT decomposes under these specific conditions. Single peaks eluting at 8.2 min were obtained for all the tested TNT concentrations. None of the samples from the tissue extracts (CYP81D11 nor empty vector) presented any peak eluting at that retention time nor with the same UV trace. However, peaks eluting after 6.5 min and at  $\sim 9.1$  min were observed. This meant that in the tissue samples the TNT taken up from the liquid medium was totally converted into two derivatives. On the average of the replicates, the CYP81D11-containing samples produced an amount of the 9.1' product approximately 1.9-times higher than the empty vector controls (figure 5.17). Nevertheless, the statistical analysis (T-test) of the data demonstrated that the observed difference is not convincing enough to consider the two data sets from the *CYP81D11*-expressors and the negative control significantly different.



**Figure 5.17:** Comparison of the peak areas of the tobacco tissue extracts. Represented are the means of 8 biological replicates  $\pm$  S.E.

In order to verify the identity of the resulting compounds, the samples were subjected to electro-spray ionisation mass spectrometry (ESI-MS), in negative mode. This analysis showed that the 6.5' peak yielded a correspondent fragment with an  $m/z$  ratio of 196, exactly as the commonly found amino-dinitrotoluene derivatives.

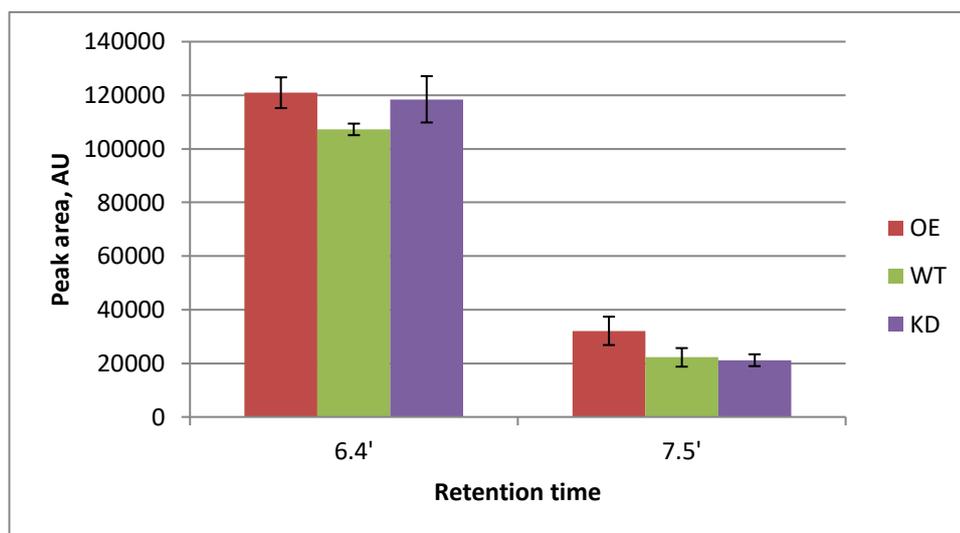
For the peak eluting at 9.1 min, a fragment was generated, with an  $m/z$  value of 405 (figure 5.18). This metabolite was observed only in the plant tissues and was not detected in the medium.



**Figure 5.18:** ESI spectra resulting from the MS analysis, in negative mode, of the metabolites extracted from the tobacco leaf discs. Prior to the MS analysis, the analytes in the extracts were separated via HPLC, using an isocratic analytical method of 50:50 water:acetonitrile. The fragments of interest, corresponding to the 6.5' (top panel) and the 9.1' (bottom panel) peaks, are circled in blue, with relative structures represented on the side.

In order to verify this outcome, a similar incubation experiment, following the same procedure as for the tobacco samples (see section 5.3.9), was carried out using the

Arabidopsis *CYP81D11*-modified lines (overexpressing and knockdown). The result of the HPLC analysis of the tissue extracts is reported in figure 5.19.



**Figure 5.19:** Comparison of the peak areas deriving from the chromatographic separation of TNT derivatives from Arabidopsis incubations, performed following the same procedure as for the tobacco transformed lines. Assay performed with eight biological replicates  $\pm$  SE.

As shown in figure 5.19, two compounds were formed during the incubation period: one eluted at 6.4 min and displayed the same UV trace as the 6.5' compound formed in the tobacco incubations. Additionally, a second peak was visualised, eluting approximately one min after the first peak, at 7.4 min. No significant difference in the formation of both the derivatives could be deduced by comparing the data sets for the three plant lines - *CYP81D11*-overexpressor, *CYP81D11*-knockdown versus wild type – used in this study.

## 5.4 Discussion

Thousands of cytochrome P450 gene sequences have been annotated to date and the most numerous group is constituted by plant CYPs. As an example, Arabidopsis alone encodes for more than 240 P450 sequences (in comparison, humans present only 57 CYPs), as revealed by genome sequencing (17).

Of the set of plant P450 enzymes studied to date, many have been attributed roles in essential mechanisms *in planta*, such as structural development, signaling and stress responses (17). Additionally, plant CYPs have also been described as key to a multitude of synthetic pathways, leading to high-value compounds for human consumption (nutraceuticals, aromas and flavours, dyes, pharmaceuticals) (109, 255). Mirroring the detoxification function of human hepatic P450s towards drugs, plant CYPs can also be active towards herbicides and pesticides found in the environment, with promising agronomical and phytoremediation applications. An area which has not been explored so far is the use of plant CYPs to help in the remediation of explosive-contaminated environments. Previous investigations carried out in our group demonstrated the potential of an array of plant enzymes, such as OPRs (163), UGTs (162) and GSTs (164, 256) for this purpose. These groups of enzymes were firstly highlighted in analyses of gene expression in Arabidopsis plants exposed to TNT and other xenobiotics (160, 162, 257). Hypothesising a direct correlation between the enhanced expression of these particular classes of enzymes with a detoxifying activity towards the target toxic compounds, previous members of Prof. Bruce's group performed direct assays with heterologously expressed and purified OPRs, UGTs and GST enzymes, as well as with transgenic plants exposed to TNT (162-164, 258).

Cytochromes P450 were also among the upregulated genes in the TNT-treated plants and their involvement in TNT detoxification has not been investigated until now. An explanation for this could be the vastness of literature describing the reductive catabolism of TNT (241), not only by plant but also by microbial enzymes; mechanisms which do not fall within the standard P450 oxidative catalysis. In addition to the expression data gathered by Dr. Lorenz's microarray study (161) and the serial analysis of gene expression by Ekman *et al.* (160), a further hint on the potential role of CYPs in TNT detoxification can be deduced by analysing the co-expression patterns of the most TNT-upregulated CYP, CYP81D11, on the CYPedia platform (159, 251). In this database, CYP81D11 is expressed in the same tissues or organs, stress situations and upon the same hormonal stimuli as the TNT-detoxifying GSTs and UGTs, thus drawing the supposition that CYP81D11 might be involved in

reactions upstream or downstream to those catalysed by the co-expressed enzymes. Furthermore, Bhadra *et al.* (1999) reported the production of a set of oxidative derivatives of TNT in liquid incubations of parrot's-feather plants with the explosive, therefore providing evidence that TNT can also undergo oxidative attack. Similar TNT derivatives have been found also by Bruns-Nagel in TNT-contaminated soils (250).

Encouraged by these preliminary observations, multiple assays were carried out, employing *CYP81D11*-overexpressing and knockdown lines of Arabidopsis plants as well as tobacco leaf discs transiently expressing CYP81D11. Both overexpressing and knockdown lines of CYP81D11 were kindly donated by Prof. J. Napier from Rothamsted Research, where CYP81D11 was the subject of studies on plant-insect multitrophic interactions and biotic stress responses.

In order to verify the respective levels of CYP81D11 expression in each line (wild type, versus overexpressing and knockdown) a qPCR analysis was performed on tissues harvested from two-week old Arabidopsis plants of all three lines. In the absence of TNT overexpressing lines displayed, as expected, elevated levels of CYP81D11 transcripts. Wild type and knockdown lines exhibited similar levels of expression, suggesting that the expression of CYP81D11 is regulated and depends on external stimuli, in this case TNT. When TNT was added in the medium, the plants showed significantly different CYP81D11 transcript levels.

In general, when the *CYP81D11*-modified plants were grown on TNT-containing agar plates and liquid media, no significant difference between wild type and overexpressing plants was observed in terms of root length and TNT uptake respectively. On the other side, interesting outcomes could be noted for the knockdown plants. A deduction that can be made, upon examination of the results obtained for the wild type and overexpressing lines, is that the overexpression of a CYP reductase partner might also be required concomitantly, to efficiently support the augmented activity, i.e. boosted electron flux, related to the overexpression of the P450 heme.

A comparison of the root length of plants grown on agar plates with a range of TNT concentrations was used as an indicator of the plants' health and resilience to the

xenobiotic. The observed outcome (at TNT concentrations of 15  $\mu\text{M}$  and 30  $\mu\text{M}$  knockdown lines displayed significantly shorter roots than wild type and overexpressing lines) could signify that, by producing lower amounts of CYP81D11, the overall defence mechanism of the plant to high concentrations of TNT was weakened, rendering the seedlings not able to grow further under such conditions. This result contradicts the results of the liquid media experiments (where knockdown plants displayed faster uptake of TNT and no visible sign of higher toxicity, if compared to the wild type and knockdown lines) and is hard to explain. Furthermore, a lot of variation was noted across the TNT concentrations and different replicates, to be able to draw a safe conclusion. Clearly this experiment needs to be repeated, possibly with lower amounts of seeds in each plate and using fresh seeds batches.

For further investigations on the function of CYP81D11 *in planta*, *N. benthamiana* leaves were *Agrobacterium*-infiltrated to transiently express the target CYP. A BlastP sequence analysis could not find matches exceeding 42% sequence identity over 28 % of the CYP81D11 sequence, which was the result found for a tobacco endogenous N-demethylase (sequence ID: ANF07088.1). This proves that the tobacco system does not present a CYP81D11 orthologue and therefore no overlapping endogenous activity should be observed. The HPLC analysis of the samples collected showed that all the supplied TNT (100  $\mu\text{M}$ ) was removed from the medium within 48 h. In the tobacco tissue extracts, two types of TNT derivatives were found, with retention times of 6.5' and 9'. The first compound (eluting at 6.5 min) presented the same UV trace and MS fragmentation pattern of the aminodinitrotoluene derivatives, typical derivatives resulting from the reduction of TNT, and probably the result of endogenous nitroreductases present in the tissues. Nitroreductases might also be responsible for the production of the second metabolite observed (eluting at 9 min), which, upon electrospray ionisation, generated an ion with an m/z ratio equal to that observed for azoxy-tetranitrotoluenes (m/z = 405). The formation of a macromolecule with such m/z ratio, resulting from the condensation of two reduced molecules of TNT, has been also described in the past after incubations with TNT of a purified rat liver reductase (259), of fungal mycelia (260, 261) and of batches of contaminated soil (with related indigenous bacteria (262)).

When extracts from *CYP81D11*-overexpressing and wild type Arabidopsis plant, from plants incubated with 100  $\mu$ M TNT (as for the tobacco experiment) for 48 h, were analysed by HPLC, two derivative molecules could be observed again, with retention times of 6.4 and 7.5 min. While the 6.4' peak corresponded to the already identified ADNTs, MS analysis is still required to identify the second product (eluting at 7.5 min). In both cases, there was no significant difference in the amount produced, between the wild type and the *CYP81D11*-overexpressing plants, suggesting that in the examined Arabidopsis lines CYP81D11 does not perform any direct action towards the given substrate (nor subsequent derivatives) in the 48 h-time frame.

Overall, the results obtained in this chapter evidenced that CYP81D11 is indeed implicated in the stress response by the plant towards TNT. This role could be further proved by carrying out metabolite profiling in the CYP81D11-knockdown Arabidopsis plant extracts and compare these result to the profiles for the wild type and overexpressing plants. This approach could also be useful to determine whether other parallel pathways intervene when CYP81D11 is downregulated. A direct activity of CYP81D11 towards TNT could be proven by performing activity assays with the purified enzyme. Moreover, it has to be taken into account that the Arabidopsis CYP81D11 transgenic plants (knockdown and overexpressing) used here are single lines donated by Prof. Napier. For this reason, such studies, for accuracy and reliability of the results, should be repeated after the generation of multiple independent *CYP81D11*-modified lines. It could also be argued that the levels of CYP81D11 transcripts in the modified lines are not sufficiently different from the wild type levels, thus the lines used in this chapter might not be optimal to be able to observe specific, measurable, phenotypes upon TNT exposure. In addition, clearer results on the effective physiological role of CYP81D11 could be obtained by depleting completely the expression of the enzyme, generating *CYP81D11*-knockout lines.

Previous studies on Arabidopsis plants' response to chemicals conducted by Baerson *et al.*, highlighted that CYP81D11 is the most upregulated CYP upon exposure to benzoxazolin-2(3H)-one (BOA), a benzoxazinoid derivative belonging to the family of allelochemicals and antimicrobial/antiherbivore defence compounds produced by

plants. In Baerson's study, numerous other chemicals induced the expression of CYP81D11, such as drugs, herbicide safeners and 2,4,5-trichlorophenol (20 mM phenobarbital, 100  $\mu$ M Fenclorim, 0.1-1 mM Benoxacor, 100  $\mu$ M TCP, respectively) (263). Additionally, members of Prof. Napier's laboratory, demonstrated that CYP81D11 in *Arabidopsis* is involved in the *cis*-jasmone-dependent plant defence to herbivore attacks. In particular, upon stimulation, *CYP81D11*-overexpressing plants mediated the production of volatile signaling compounds, affecting the interactions between insects (and their pests) and plants (264-266). Recently, Walper *et al.* reported that CYP81D11 expression is triggered in response to oxylipins (specifically to 9-hydroxy-10,12,15-octadecatrienoic acid, 9-HOX). In the test carried out in Walper's study, CYP81D11 overexpressor and wild type plants were grown on medium containing 9-HOX. The plants overexpressing the CYP displayed longer roots, i.e. higher tolerance to the chemical (267, 268).

All these findings lead us to the consensus view that CYP81D11 might not be directly active towards TNT specifically, but instead it could be part of a general defence response mechanism, induced by external biotic and abiotic stresses.

## Chapter 6: Final discussion

Cytochromes P450 are heme-thiolate monooxygenases that can be found in almost all organisms, from microbes and mammals to insects and plants. CYPs catalyse important activities, from the synthesis of essential endobiotics (such as vitamins, sterols, hormones, fatty acids) to the metabolism of xenobiotic compounds (such as drugs and environmental pollutants). Since their discovery, in 1958, and the first characterisation studies by Omura and Sato a few years later, cytochromes P450 have attracted the attention of numerous research groups around the world. One of the reasons for the significant interest in CYPs is the potential use of these enzymes as biocatalysts for biotechnological applications, such as engineered processes for the production of high-value chemicals.

The pharmaceutical industry has particularly benefited from P450-related research: a wide array of bacterial, plant and fungal CYPs have been discovered in the synthetic pathways leading to chemotherapeutics, antimicrobial, antioxidant, antimalarial compounds (15, 106, 114, 221, 223, 255, 269-271). Additionally, the mammalian hepatic CYPs have contributed extensive data in the context of drug metabolism; data of particular importance for pharmacokinetics and toxicity studies during drug development.

In plants, given the role of cytochromes P450 not only in the metabolism of exogenous compounds, but also in endogenous signalling and structural development pathways, CYPs can be engineered to enhance tolerance to biotic and abiotic stresses (180, 192, 272), to extract pollutants from the environment (23, 235) or to enhance the nutritional value of crops (125).

Of the 35,000+ CYP sequences (173), most have been found in plants, subdivided in 127 families and, on the basis of phylogeny, eleven clans (101). Besides the high sequence divergence found between the CYPs, another constraint in the functional and structural characterisation of plant cytochromes P450, is the natural insolubility of these enzymes, due to the hydrophobic N-terminal domain and internal hydrophobic patches. Whilst in contrast, microbial CYPs are generally soluble enzymes, therefore, easier to isolate, reconstitute and characterise.

In this project a genetic fusion construct was developed, to ease the heterologous expression and functional characterisation of plant CYPs with unknown functions. In addition, a variety of *in vitro* and *in vivo* assays were carried out in order to evaluate the potential role of Arabidopsis CYP81D11 in the detoxification of the explosive TNT.

Chapter 3 focused on the screening for optimal conditions for the soluble expression of plant CYPs in *E. coli*. The target cytochromes P450 were the well-described cinnamate hydroxylase (CYP73A5) from Arabidopsis, CYP81D11 from Arabidopsis as well, and maize CYP81A9. The first target, CYP73A5, was chosen for the availability of the standards for activity assays, whereas the other two were selected due to their potential application in the detoxification of xenobiotics (in particular explosives for CYP81D11 (160, 162) and herbicides for CYP81A9 (168), on the basis of their expression patterns). The replacement of the N-terminal membrane anchor with a short string of amino acids (sequence: MAKKTSSKG, named in chapter 3 'S'-peptide), the codon-optimisation of the sequences for expression in *E. coli* and the use of the Arctic express strain as host for expression, enabled the production of all three CYP targets, as soluble products. This successful outcome was demonstrated with the visualisation of bands with the correct size on SDS-PAGE gels and by western blot analysis. No CO-differential spectrum could be recorded for any of the purified CYP targets, showing that the replacement of the natural N-terminal hydrophobic domain, and possibly also the chosen purification conditions, have affected the structural configuration of the proteins. Activity assays performed with the purified S-CYP73A5tr did not yield the expected product, indicating that the protein was not active, probably due to unfolding. Additionally, the assay conditions (such as temperature and proportion of CYP and supplied reductase) might not have been optimal to be able to observe catalysis. As an alternative, plasmids carrying the native sequences of AtCYP81D11 and AtCYP81D8, were transformed into yeast WAT11 cells (co-expressing the AtATR1) for expression, with no success.

In Chapter 4, a fusion construct for the co-expression of CYPs with a suitable redox partner was built. The first component of the platform was a cytochrome P450 reductase from Arabidopsis, ATR2, which had been previously characterised and compared in terms of catalytic efficiency to ATR1, the other CYP reductase found in

Arabidopsis, by Dr. Schuckel (152). The chosen reductase was truncated to remove the N-terminal hydrophobic membrane-binding domain and codon-optimised for expression in *E. coli*. Subsequently, a short linker (sequence: GSTSSGSG, from Dr. Schuckel's ACRyLAM construct (211) and from Koffas' fusion of isoflavone synthase with a plant reductase (193)) was introduced upstream of the reductase, as a spacer region to enable the appropriate spatial arrangement of the heme and reductase domains. The well-characterised Arabidopsis CYP73A5 was chosen as a reference to validate the fusion construct. The CYP was modified in the same way as the reductase partner and cloned via Infusion upstream of the poly-GlySer linker. In addition, multiple N-terminal tags were introduced, in order to increase the solubility of the product and to ease the subsequent purification process: the 'S' peptide (from chapter 3), the conventional His, MBP, GST, GFP tags and the Im9 colicin immunity protein. Although it had been effective for the soluble expression of CYP73A5, CYP81D11 and CYP81A9, the 'S' tag did not have a beneficial effect in the expression of the CYP-reductase fusion. Among the other tags, Im9 allowed the production of the chimera. Purification trials were unsuccessful, but activity assays performed with whole *E. coli* cells expressing the Im9-CYP73A5tr-ATR2tr construct showed that the catalyst was active towards the given substrate, cinnamic acid, converting 60 % of the compound into the hydroxylated derivative, *p*-coumaric acid, in six h. This assay was performed with exactly the same conditions used in the past by Dr. Schuckel for the evaluation of the ACRyLAM and ACRyLIC fusion constructs, with CYP73A5; in that case the maximum conversion (40 %) was achieved with the ACRyLAM construct, in 48 h (211).

In Chapter 5, assays with modified Arabidopsis and tobacco plants were carried out, using a set of established protocols from Prof. Bruce's and Prof. Werck's groups.

With the aim of identifying novel TNT-detoxifying enzymes for phytoremediation purposes, Dr. Lorenz (a past member of Prof. Bruce's group), conducted microarray gene expression analyses in Arabidopsis tissues following exposure to the explosive TNT (161, 162). The results obtained confirmed the outcomes of another genetic analysis published by Ekman *et al.* (160). In both studies, among the most highly induced transcripts by the TNT treatment were GSTs, UGTs, CYPs, transporters,

signalling proteins and transcription factors. Therefore, research has focused on the characterisation of those highly induced groups of enzymes. CYPedia co-expression analyses data (co-localising P450s with known xenobiotic metabolisers (159)) as well as the identification of oxidative derivatives of TNT (165, 250), motivated the study of the potential role of the upregulated CYPs in the detoxification of TNT, upstream of the GST and UGT activities.

In Dr. Lorenz's study, CYP81D11 was revealed as the most upregulated CYP (24 times, compared to the untreated) in Arabidopsis, in response to TNT exposure.

Transgenic *CYP81D11*-overexpressing and *CYP81D11*-knockdown Arabidopsis plants, kindly donated by Prof. Napier (Rothamsted Research), were employed to verify whether CYP81D11 confers tolerance to TNT.

Three-week old Arabidopsis seedlings were grown on TNT-containing liquid medium and TNT removal in a time frame of 168 h (one week) was analysed via HPLC. The knockdown lines were able to take up the compound present in the medium significantly faster ( $p > 0.05$ ) than the wild type and the overexpressing lines. Specifically, more than 90 % of TNT (from an initial amount of 200  $\mu\text{M}$ ) was removed in the first 48 h, whereas the other two lines were able to reach the same level only after one week. No difference could be observed in terms of toxicity signs (chlorosis, reduced shoots/roots size) across the three lines.

When the three Arabidopsis lines were germinated and grown on TNT-containing agar plates, the CYP81D11-knockdown plants displayed lower tolerance towards TNT than the wild type and overexpressing lines, showing significantly shorter roots (-38.1 % in plates with 15  $\mu\text{M}$  TNT and -40.7 % in plates with 30  $\mu\text{M}$  TNT) after 20 days of growth.

The liquid incubation study showed that lower levels of CYP81D11 (in the *CYP81D11*-knockdown line) lead to an increase of the removal rate of TNT from the medium, with no signs of increased toxicity. In the root length experiment, knockdown plants exhibited lower tolerance to TNT. The contrasting results between the liquid and agar incubations might have been influenced by the different developmental stage in

which the plants have been assayed. The plants were germinated on the TNT-containing agar, whereas in the liquid medium experiment three-week old plants were incubated with TNT. In addition, there only one independently-transformed plant line for the *CYP81D11*-overexpressor and *CYP81D11*-knockdown Arabidopsis plants was available for the assays. Other independent lines need to be generated, and the assays have to be repeated with a suitable number of replicates, in order to minimise the experimental variation and errors, as well as to validate the outcomes.

To investigate the possible TNT derivatives coming from CYP81D11 activity, tobacco leaves were infiltrated with Agrobacteria expressing CYP81D11. The HPLC analysis of the TNT-containing medium, from the leaf discs incubations, showed that all the TNT (initial concentration: 100  $\mu$ M) was removed in 48 h. The HPLC-MS analysis of the tissue extracts demonstrated that mainly ADNTs and tetranitroazoxy-derivatives were formed, possibly deriving from endogenous nitroreductases and from the condensation of the reduced aromatic rings, as described in the literature (262).

## **Future perspectives**

### **Fusion platform**

Given the success with AtCYP73A5, the Im9 fusion platform here developed could be useful for the expression of other unknown plant cytochromes P450. The C-terminal Arabidopsis ATR2 should enable an efficient electron flux from NADPH to the heme. The strategically-designed flanking restriction sites could be exploited for the relatively easy InFusion cloning of different CYP targets in the platform, as well as, (eventually, for comparative purposes) for the introduction of other reductase partners. In addition, the 48-well screening platform used in this study could be of great help for the parallel screening of multiple expression conditions, to tailor the process to the specific CYP/CYP-reductase targets, reducing operational time and costs.

### **Elucidation of the role of CYP81D11 in TNT detoxification**

The experimental work carried out in this project did not demonstrate activity of CYP81D11 towards TNT. This outcome requires further investigation and additional experiments, which could not be carried out due to time constraints. The work that needs to be performed consists in the repetition of the TNT uptake liquid medium experiments and root length comparisons with a larger pool of independent transgenic lines, including *CYP81D11*-knockout plants in place of the knockdown lines. In addition, heterologously expressed and purified CYP81D11, should be assayed directly towards TNT or its derivatives.

# Abbreviations

<b>1/2 MS</b>	Murashige and Skoog medium, half strength
<b>ADNT</b>	Aminodinitrotoluene
<b>ADP</b>	Adenosine diphosphate
<b>ADR</b>	Adrenodoxin Reductase
<b>ADX</b>	Adrenodoxin
<b>AI</b>	Autoinduction medium
<b>ATR1</b>	Arabidopsis thaliana P450 reductase 1
<b>ATR2</b>	Arabidopsis thaliana P450 reductase 2
<b>BSA</b>	Bovine serum albumin
<b>cDNA</b>	Complementary DNA
<b>CO</b>	Carbon monoxide
<b>CPR</b>	Cytochrome P450 oxidoreductase
<b>CYP</b>	Cytochrome P450
<b>Cyt<sub>b5</sub></b>	Cytochrome b5
<b>DNA</b>	Deoxyribonucleic acid
<b>DNTP</b>	Deoxyribonucleotide triphosphate
<b>DTT</b>	Dithiothreitol
<b>EDTA</b>	Ethylenediaminetetraacetic acid
<b>ER</b>	Endoplasmic reticulum
<b>EV</b>	Empty vector
<b>FAD</b>	Flavin adenine dinucleotide
<b>FMN</b>	Flavin mononucleotide
<b>GFP</b>	Green Fluorescent Protein
<b>GST</b>	Glutathione S-transferase
<b>h</b>	Hour(s)
<b>HADNT</b>	4-Hydroxyl-2,6-dinitrotoluene
<b>His</b>	Histidine
<b>HPLC</b>	High-performance liquid chromatography

<b>IM9</b>	Colicin E9 immunity protein
<b>IPTG</b>	Isopropyl $\beta$ -D-1-thiogalactopyranoside
<b>KD</b>	Knockdown
<b>LB</b>	Luria-Bertani medium
<b>MBP</b>	Maltose-binding protein
<b>min</b>	Minutes
<b>MS</b>	Mass spectrometry
<b>NADH</b>	Reduced nicotinamide adenine dinucleotide
<b>NADP</b>	Nicotinamide adenine dinucleotide phosphate
<b>NADPH</b>	Reduced nicotinamide adenine dinucleotide phosphate
<b>NODT</b>	4-Nitroso-2,6-dinitrotoluene
<b>NPC</b>	No plant control
<b>OD</b>	Optical density
<b>OE</b>	Overexpressing
<b>OPR</b>	Oxo-phytodienoate reductase
<b>P450</b>	Cytochrome P450 enzyme
<b>PCR</b>	Polymerase chain reaction
<b>RDX</b>	Hexahydro-1,3,5-trinitro-1,3,5-triazine
<b>RNA</b>	Ribonucleic acid
<b>rpm</b>	Revolutions per minute
<b>s</b>	Seconds
<b>SDS-PAGE</b>	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
<b>SE</b>	Standard error of the mean
<b>TB</b>	Terrific broth
<b>TBAP</b>	Tetrabutylammonium dihydrogen phosphate
<b>TCE</b>	Trichloroethylene
<b>TEMED</b>	Tetramethylethylenediamine
<b>TNT</b>	2,4,6-trinitrotoluene
<b>UGT</b>	Uridine diphosphate glycosyltransferase
<b>v/v</b>	Volume to volume ratio
<b>w/v</b>	Weight to volume ratio
<b>WT</b>	Wild type

## References

1. Strittmatter CF, Ball EG. A Hemochromogen Component of Liver Microsomes. *Proc Natl Acad Sci USA*. 1952;38(1):19-25.
2. Strittmatter P, Velick SF. The isolation and properties of microsomal cytochrome. *J Biol Chem*. 1956;221(1):253-64.
3. Garfinkel D. Isolation and properties of cytochrome b5 from pig liver. *Arch Biochem Biophys*. 1957;71(1):111-20.
4. Klingenberg M. Pigments of rat liver microsomes. *Arch Biochem Biophys*. 1958;75(2):376-86.
5. Garfinkel D. Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. *Arch Biochem Biophys*. 1958;77(2):493-09.
6. Omura T, Sato R. A new cytochrome in liver microsomes. *J Biol Chem*. 1962;237:1375-76.
7. Omura T, Sato R. The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J Biol Chem*. 1964;239:2370-78.
8. Omura T, Sato R. Fractional solubilization of haemoproteins and partial purification of carbon monoxide-binding cytochrome from liver microsomes. *Biochim Biophys Acta*. 1963;71:224-26.
9. Omura T, Sato R. The carbon monoxide-binding pigment of liver microsomes. II. Solubilization, purification and properties. *J Biol Chem*. 1964;239:2379-85.
10. Ryan KJ, Engel LL. Hydroxylation of steroids at carbon 21. *J Biol Chem*. 1957;225(1):103-14.
11. Cooper DY, Levin S, Narasimhulu S, *et al*. Photochemical action spectrum of the terminal oxidase of mixed function oxidase systems. *Science*. 1965;147(3656):400-02.
12. Katagiri M, Ganguli BN, Gunsalus IC. A soluble cytochrome P-450 functional in methylene hydroxylation. *J Biol Chem*. 1968;243(12):3543-46.
13. Hedegaard J, Gunsalus IC. Mixed function oxidation. IV. An induced methylene hydroxylase in camphor oxidation. *J Biol Chem*. 1965;240(10):4038-43.
14. Appleby AC. A soluble haemoprotein P 450 from nitrogen-fixing Rhizobium bacteroids. *Biochim Biophys Acta*. 1967;147(2):399-02.

15. Werck-Reichhart D, Feyereisen R. Cytochromes P450: a success story. *Genome Biol.* 2000;1(6):Reviews3003.1–3.9.
16. Nelson DR. Cytochrome P450 nomenclature, 2004. *Methods Mol Biol.* 2006;320:1-10.
17. Bak S, Beisson F, Bishop G, *et al.* Cytochromes p450. *Arabidopsis Book.* 2011;9:e0144.
18. Poulos TL, Finzel BC, Howard AJ. High-resolution crystal structure of cytochrome P450cam. *J Mol Biol.* 1987;195(3):687-700.
19. Buchholz P, Vogel C, Reusch W, *et al.* BioCatNet: a database system for the integration of enzyme sequences and biocatalytic experiments. *Chembiochem.* 2016;17(21):2093-98.
20. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature.* 2000;408(6814):796-15.
21. Nelson DR, Schuler MA, Paquette SM, *et al.* Comparative genomics of rice and Arabidopsis. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot. *Plant Physiol.* 2004;135(2):756-72.
22. Hannemann F, Bichet A, Ewen KM, *et al.* Cytochrome P450 systems--biological variations of electron transport chains. *Biochim Biophys Acta.* 2007;1770(3):330-44.
23. Rylott EL, Jackson RG, Sabbadin F, *et al.* The explosive-degrading cytochrome P450 XplA: biochemistry, structural features and prospects for bioremediation. *Biochim Biophys Acta.* 2011;1814(1):230-36.
24. Joo H, Lin Z, Arnold FH. Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. *Nature.* 1999;399(6737):670-73.
25. Zhang H, Im SC, Waskell L. 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.* 2007;282(41):29766-76.
26. Vermilion JL, Coon MJ. Identification of the high and low potential flavins of liver microsomal NADPH-cytochrome P-450 reductase. *J Biol Chem.* 1978;253(24):8812-19.
27. Wang M, Roberts DL, Paschke R, *et al.* Three-dimensional structure of NADPH-cytochrome P450 reductase: prototype for FMN- and FAD-containing enzymes. *Proc Natl Acad Sci USA.* 1997;94(16):8411-16.

28. Sündermann A, Oostenbrink C. Molecular dynamics simulations give insight into the conformational change, complex formation, and electron transfer pathway for cytochrome P450 reductase. *Protein Sci.* 2013;22(9):1183-95.
29. Iyanagi T, Xia C, Kim JJ. NADPH-cytochrome P450 oxidoreductase: prototypic member of the diflavin reductase family. *Arch Biochem Biophys.* 2012;528(1):72-89.
30. Shephard EA, Phillips IR, Bayney RM, *et al.* Quantification of NADPH: cytochrome P-450 reductase in liver microsomes by a specific radioimmunoassay technique. *Biochem J.* 1983;211(2):333-40.
31. Xia C, Panda SP, Marohnic CC, *et al.* Structural basis for human NADPH-cytochrome P450 oxidoreductase deficiency. *Proc Natl Acad Sci USA.* 2011;108(33):13486-91.
32. Lu AY, Coon MJ. Role of hemoprotein P-450 in fatty acid omega-hydroxylation in a soluble enzyme system from liver microsomes. *J Biol Chem.* 1968;243(6):1331-32.
33. Porter TD. New insights into the role of cytochrome P450 reductase (POR) in microsomal redox biology. *Acta Pharm Sin B.* 2012;2(2):102-06.
34. McLean KJ, Sabri M, Marshall KR, *et al.* Biodiversity of cytochrome P450 redox systems. *Biochem Soc Trans.* 2005;33(Pt 4):796-01.
35. Pikuleva IA, Tesh K, Waterman MR, *et al.* The tertiary structure of full-length bovine adrenodoxin suggests functional dimers. *Arch Biochem Biophys.* 2000;373(1):44-55.
36. Beilke D, Weiss R, Löhr F, *et al.* A new electron transport mechanism in mitochondrial steroid hydroxylase systems based on structural changes upon the reduction of adrenodoxin. *Biochemistry.* 2002;41(25):7969-78.
37. Ewen KM, Kleser M, Bernhardt R. Adrenodoxin: The archetype of vertebrate-type [2Fe-2S] cluster ferredoxins. *Biochim Biophys Acta, Proteins Proteomics.* 2011;1814(1):111-25.
38. Schiffler B, Bernhardt R. Bacterial (CYP101) and mitochondrial P450 systems-how comparable are they? *Biochem Biophys Res Commun.* 2003;312(1):223-28.
39. Neunzig I, Widjaja M, Peters FT, *et al.* Coexpression of CPR from various origins enhances biotransformation activity of human CYPs in *S. pombe*. *Appl Biochem Biotechnol.* 2013;170(7):1751-66.

40. Schiffler B, Kiefer M, Wilken A, *et al.* The interaction of bovine adrenodoxin with CYP11A1 (cytochrome P450<sub>scc</sub>) and CYP11B1 (cytochrome P450<sub>11beta</sub>). Acceleration of reduction and substrate conversion by site-directed mutagenesis of adrenodoxin. *J Biol Chem.* 2001;276(39):36225-32.
41. Sono M, Roach MP, Coulter ED, *et al.* Heme-containing oxygenases. *Chem Rev.* 1996;96(7):2841-88.
42. Isin EM, Guengerich FP. Complex reactions catalyzed by cytochrome P450 enzymes. *Biochim Biophys Acta.* 2007;1770(3):314-29.
43. Grinkova YV, Denisov IG, McLean MA, *et al.* Oxidase uncoupling in heme monooxygenases: human cytochrome P450 CYP3A4 in nanodiscs. *Biochem Biophys Res Commun.* 2013;430(4):1223-27.
44. Krest CM, Onderko EL, Yosca TH, *et al.* Reactive intermediates in cytochrome p450 catalysis. *J Biol Chem.* 2013;288(24):17074-81.
45. Lander ES, Linton LM, Birren B, *et al.* Initial sequencing and analysis of the human genome. *Nature.* 2001;409(6822):860-21.
46. Nebert DW, Wikvall K, Miller WL. Human cytochromes P450 in health and disease. *Philos Trans R Soc Lond B Biol Sci.* 2013;368(1612):20120431.
47. Thelen K, Dressman JB. Cytochrome P450-mediated metabolism in the human gut wall. *J Pharm Pharmacol.* 2009;61(5):541-58.
48. Ferguson CS, Tyndale RF. Cytochromes P450 in the brain: Emerging evidence for biological significance. *Trends Pharmacol Sci.* 2011;32(12):708-14.
49. Simpson ER, Mahendroo MS, Means GD, *et al.* Aromatase cytochrome P450, the enzyme responsible for estrogen biosynthesis. *Endocr Rev.* 1994;15(3):342-55.
50. Bouchoucha N, Samara-Boustani D, Pandey AV, *et al.* Characterization of a novel CYP19A1 (aromatase) R192H mutation causing virilization of a 46,XX newborn, undervirilization of the 46,XY brother, but no virilization of the mother during pregnancies. *Mol Cell Endocrinol.* 2014;390(1-2):8-17.
51. Daldorff S, Mathiesen RM, Yri OE, *et al.* Cotargeting of CYP-19 (aromatase) and emerging, pivotal signalling pathways in metastatic breast cancer. *Br J Cancer.* 2017; 116(1): 10–20.
52. Berstein LM, Imyanitov EN, Kovalevskij AJ, *et al.* CYP17 and CYP19 genetic polymorphisms in endometrial cancer: association with intratumoral aromatase activity. *Cancer Lett.* 2004;207(2):191-96.

53. Guengerich FP. Intersection of the roles of cytochrome P450 enzymes with xenobiotic and endogenous substrates: relevance to toxicity and drug interactions. *Chem Res Toxicol*. 2017;30(1):2-12.
54. Williams JA, Hyland R, Jones BC, *et al*. Drug-drug interactions for UDP-glucuronosyltransferase substrates: a pharmacokinetic explanation for typically observed low exposure (AUC<sub>i</sub>/AUC) ratios. *Drug Metab Dispos*. 2004;32(11):1201-08.
55. Stiborova M, Sejbal J, Borek-Dohalska L, *et al*. The anticancer drug ellipticine forms covalent DNA adducts, mediated by human cytochromes P450, through metabolism to 13-hydroxyellipticine and ellipticine N2-oxide. *Cancer Res*. 2004;64(22):8374-80.
56. Roy P, Yu LJ, Crespi CL, *et al*. Development of a substrate-activity based approach to identify the major human liver P-450 catalysts of cyclophosphamide and ifosfamide activation based on cDNA-expressed activities and liver microsomal P-450 profiles. *Drug Metab Dispos*. 1999;27(6):655-66.
57. Gallagher EP, Wienkers LC, Stapleton PL, *et al*. Role of human microsomal and human complementary DNA-expressed cytochromes P4501A2 and P4503A4 in the bioactivation of aflatoxin B1. *Cancer Res*. 1994;54(1):101-08.
58. Yun CH, Shimada T, Guengerich FP. Roles of human liver cytochrome P4502C and 3A enzymes in the 3-hydroxylation of benzo(a)pyrene. *Cancer Res*. 1992;52(7):1868-74.
59. Kawahigashi H, Hirose S, Ohkawa H, *et al*. Phytoremediation of the herbicides atrazine and metolachlor by transgenic rice plants expressing human CYP1A1, CYP2B6, and CYP2C19. *J Agric Food Chem*. 2006;54(8):2985-91.
60. Kawahigashi H, Hirose S, Ohkawa H, *et al*. Transgenic rice plants expressing human p450 genes involved in xenobiotic metabolism for phytoremediation. *J Mol Microbiol Biotechnol*. 2008;15(2-3):212-19.
61. Shiota N, Kodama S, Inui H, Ohkawa H. Expression of human cytochromes P450 1A1 and P450 1A2 as fused enzymes with yeast NADPH-cytochrome P450 oxidoreductase in transgenic tobacco plants. *Biosci Biotechnol Biochem*. 2000;64(10):2025-33.
62. James CA, Xin G, Doty SL, Strand SE. Degradation of low molecular weight volatile organic compounds by plants genetically modified with mammalian cytochrome P450 2E1. *Environ Sci Technol*. 2008;42(1):289-93.
63. Guengerich FP. Cytochrome P-450 3A4: regulation and role in drug metabolism. *Annu Rev Pharmacol Toxicol*. 1999;39:1-17.

64. Bodin K, Bretillon L, Aden Y, *et al.* Antiepileptic drugs increase plasma levels of 4beta-hydroxycholesterol in humans: evidence for involvement of cytochrome p450 3A4. *J Biol Chem.* 2001;276(42):38685-89.
65. Niwa T, Murayama N, Imagawa Y, *et al.* Regioselective hydroxylation of steroid hormones by human cytochromes P450. *Drug Metab Rev.* 2015;47(2):89-10.
66. Niwa T, Yabusaki Y, Honma K, *et al.* Contribution of human hepatic cytochrome P450 isoforms to regioselective hydroxylation of steroid hormones. *Xenobiotica.* 1998;28(6):539-47.
67. Waxman DJ, Attisano C, Guengerich FP, *et al.* Human liver microsomal steroid metabolism: identification of the major microsomal steroid hormone 6 beta-hydroxylase cytochrome P-450 enzyme. *Arch Biochem Biophys.* 1988;263(2):424-36.
68. Inoue K, Sodhi K, Puri N, *et al.* Endothelial-specific CYP4A2 overexpression leads to renal injury and hypertension via increased production of 20-HETE. *Am J Physiol - Renal.* 2009;297(4):F875-84.
69. Wu CC, Mei S, Cheng J, *et al.* Androgen-sensitive hypertension associates with upregulated vascular CYP4A12-20-HETE synthase. *J Am Soc Nephrol.* 2013;24(8):1288-96.
70. Savas U, Wei S, Hsu MH, *et al.* 20-Hydroxyeicosatetraenoic acid (HETE)-dependent hypertension in human cytochrome P450 (CYP) 4A11. Transgenic mice: normalization of blood pressure by sodium restriction, hydrochlorothiazide, or blockade of the type I angiotensin II receptor. *J Biol Chem.* 2016;291(32):16904-19.
71. Sim SC, Ingelman-Sundberg M. The Human Cytochrome P450 (CYP) Allele nomenclature website: a peer-reviewed database of CYP variants and their associated effects. *Hum Genomics.* 2010;4(4):278-81.
72. Sim SC. The Human Cytochrome P450 (CYP) Allele nomenclature database - CYP2D6 2010 [homepage on the internet] Available from: <http://www.cypalleles.ki.se/cyp2d6.htm>.
73. Mann A, Tyndale RF. Cytochrome P450 2D6 enzyme neuroprotects against 1-methyl-4-phenylpyridinium toxicity in SH-SY5Y neuronal cells. *Eur J Neurosci.* 2010;31(7):1185-93.
74. McCann SJ, Pond SM, James KM, *et al.* The association between polymorphisms in the cytochrome P-450 2D6 gene and Parkinson's disease: a case-control study and meta-analysis. *J Neurol Sci.* 1997;153(1):50-53.
75. Mann A, Miksys SL, Gaedigk A, *et al.* The neuroprotective enzyme CYP2D6 increases in the brain with age and is lower in Parkinson's disease patients. *Neurobiol Aging.* 2012;33(9):2160-71.

76. Dong Y, Xiao H, Wang Q, *et al.* Analysis of genetic variations in CYP2C9, CYP2C19, CYP2D6 and CYP3A5 genes using oligonucleotide microarray. *Int J Clin Exp Med.* 2015;8(10):18917-26.
77. von Bahr C, Spina E, Birgersson C, *et al.* Inhibition of desmethylimipramine 2-hydroxylation by drugs in human liver microsomes. *Biochem Pharmacol.* 1985;34(14):2501-05.
78. von Moltke LL, Greenblatt DJ, Court MH, *et al.* Inhibition of alprazolam and desipramine hydroxylation in vitro by paroxetine and fluvoxamine: comparison with other selective serotonin reuptake inhibitor antidepressants. *J Clin Psychopharmacol.* 1995;15(2):125-31.
79. von Moltke LL, Greenblatt DJ, Schmider J, *et al.* In vitro approaches to predicting drug interactions in vivo. *Biochem Pharmacol.* 1998;55(2):113-22.
80. Gomez DY, Wachter VJ, Tomlanovich SJ, *et al.* The effects of ketoconazole on the intestinal metabolism and bioavailability of cyclosporine. *Clin Pharmacol Ther.* 1995;58(1):15-19.
81. Koudriakova T, Iatsimirskaia E, Utkin I, *et al.* Metabolism of the human immunodeficiency virus protease inhibitors indinavir and ritonavir by human intestinal microsomes and expressed cytochrome P4503A4/3A5: mechanism-based inactivation of cytochrome P4503A by ritonavir. *Drug Metab Dispos.* 1998;26(6):552-61.
82. Kelly SL, Kelly DE. Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? *Philos Trans R Soc Lond B Biol Sci.* 2013;368(1612):20120476.
83. Narhi LO, Fulco AJ. Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P-450 monooxygenase induced by barbiturates in *Bacillus megaterium*. *J Biol Chem.* 1986;261(16):7160-69.
84. Miles JS, Munro AW, Rospendowski BN, *et al.* Domains of the catalytically self-sufficient cytochrome P-450 BM-3. Genetic construction, overexpression, purification and spectroscopic characterization. *Biochem J.* 1992;288 (Pt 2):503-09.
85. Munro AW, Girvan HM, McLean KJ. Cytochrome P450–redox partner fusion enzymes. *BBA - Gen Subjects.* 2007;1770(3):345-59.
86. Sevrioukova IF, Li H, Zhang H, *et al.* Structure of a cytochrome P450–redox partner electron-transfer complex. *P Natl Acad Sci USA.* 1999;96(5):1863-68.
87. Seth-Smith HM, Rosser SJ, Basran A, *et al.* Cloning, sequencing, and characterization of the hexahydro-1,3,5-Trinitro-1,3,5-triazine degradation gene cluster from *Rhodococcus rhodochrous*. *Appl Environ Microbiol.* 2002;68(10):4764-71.

88. Rylott EL, Jackson RG, Edwards J, *et al.* An explosive-degrading cytochrome P450 activity and its targeted application for the phytoremediation of RDX. *Nat Biotechnol.* 2006;24(2):216-19.
89. Roberts GA, Grogan G, Greter A, *et al.* Identification of a new class of cytochrome P450 from a *Rhodococcus* sp. *J Bacteriol.* 2002;184(14):3898-08.
90. Sabbadin F, Hyde R, Robin A, *et al.* LICRED: a versatile drop-in vector for rapid generation of redox-self-sufficient cytochrome P450s. *Chembiochem.* 2010;11(7):987-94.
91. Sabbadin F. Engineering cytochrome P450s for biocatalysis and bioremediation. 2010, University of York: PhD thesis.
92. Zhang W, Liu Y, Yan J, *et al.* New reactions and products resulting from alternative interactions between the P450 enzyme and redox partners. *J Am Chem Soc.* 2014;136(9):3640-46.
93. Makino T, Otomatsu T, Shindo K, *et al.* Biocatalytic synthesis of flavones and hydroxyl-small molecules by recombinant *Escherichia coli* cells expressing the cyanobacterial CYP110E1 gene. *Microb Cell Fact.* 2012;11:95.
94. Zehentgruber D, Hannemann F, Bleif S, *et al.* Towards preparative scale steroid hydroxylation with cytochrome P450 monooxygenase CYP106A2. *Chembiochem.* 2010;11(5):713-21.
95. Schmitz D, Zapp J, Bernhardt R. Steroid conversion with CYP106A2 - production of pharmaceutically interesting DHEA metabolites. *Microb Cell Fact.* 2014;13:81.
96. Bleif S, Hannemann F, Zapp J, *et al.* A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto-beta-boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Appl Microbiol Biotechnol.* 2012;93(3):1135-46.
97. Brill E, Hannemann F, Zapp J, *et al.* A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto-beta-boswellic acid (KBA). *Appl Microbiol Biotechnol.* 2014;98(4):1701-17.
98. Kiss FM, Schmitz D, Zapp J, *et al.* Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium* - identification of a novel 11-oxidase activity. *Appl Microbiol Biotechnol.* 2015;99(20):8495-514.
99. Bracco P, Janssen DB, Schallmeyer A. Selective steroid oxyfunctionalisation by CYP154C5, a bacterial cytochrome P450. *Microb Cell Fact.* 2013;12:95.

100. Peters MW, Meinhold P, Glieder A, *et al.* Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3. *J Am Chem Soc.* 2003;125(44):13442-50.
101. Hamberger B, Bak S. Plant P450s as versatile drivers for evolution of species-specific chemical diversity. *Philos Trans R Soc Lond B Biol Sci.* 2013;368(1612):20120426.
102. O'Keefe D P, Leto KJ. Cytochrome P-450 from the Mesocarp of Avocado (*Persea americana*). *Plant Physiol.* 1989;89(4):1141-49.
103. Tu Y. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nat Med.* 2011;17(10):1217-20.
104. Graham IA, Besser K, Blumer S, *et al.* The genetic map of *Artemisia annua* L. identifies loci affecting yield of the antimalarial drug artemisinin. *Science.* 2010;327(5963):328-31.
105. Artemisia research project C. Artemisia F1 seed [homepage on the Internet] Available from: <http://www.artemisiaf1seed.org/>.
106. Paddon CJ, Westfall PJ, Pitera DJ, *et al.* High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature.* 2013;496(7446):528-32.
107. Paddon CJ, Keasling JD. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat Rev Microbiol.* 2014;12(5):355-67.
108. Ro DK, Paradise EM, Ouellet M, *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature.* 2006;440(7086):940-43.
109. Renault H, Bassard JE, Hamberger B, *et al.* Cytochrome P450-mediated metabolic engineering: current progress and future challenges. *Curr Opin Plant Biol.* 2014;19:27-34.
110. WHO. Essential Medicines and Health Products 2015 [homepage on the Internet] Available from: [http://www.who.int/medicines/publications/essentialmedicines/EML\\_2015\\_FINAL\\_amended\\_NOV2015.pdf?ua=1](http://www.who.int/medicines/publications/essentialmedicines/EML_2015_FINAL_amended_NOV2015.pdf?ua=1).
111. Winkler RG, Helentjaris T. The maize Dwarf3 gene encodes a cytochrome P450-mediated early step in Gibberellin biosynthesis. *Plant Cell.* 1995;7(8):1307-17.
112. Schoendorf A, Rithner CD, Williams RM, *et al.* Molecular cloning of a cytochrome P450 taxane 10 beta-hydroxylase cDNA from *Taxus* and functional expression in yeast. *Proc Natl Acad Sci USA.* 2001;98(4):1501-06.

113. Jennewein S, Wildung MR, Chau M, *et al.* Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc Natl Acad Sci USA*. 2004;101(24):9149-54.
114. Kaspera R, Croteau R. Cytochrome P450 oxygenases of Taxol biosynthesis. *Phytochemistry rev*. 2006;5(2-3):433-44.
115. Croteau R, Ketchum REB, Long RM, *et al.* Taxol biosynthesis and molecular genetics. *Phytochemistry rev*. 2006;5(1):75-97.
116. Jensen K, Jensen PE, Moller BL. Light-driven cytochrome p450 hydroxylations. *ACS Chem Biol*. 2011;6(6):533-39.
117. Lassen LM, Nielsen AZ, Olsen CE, *et al.* Anchoring a plant cytochrome P450 via PsaM to the thylakoids in *Synechococcus* sp. PCC 7002: evidence for light-driven biosynthesis. *PLoS One*. 2014;9(7):e102184.
118. Halkier BA, Moller BL. Involvement of cCytochrome P-450 in the biosynthesis of dhurrin in *Sorghum bicolor* (L.) Moench. *Plant Physiol*. 1991;96(1):10-17.
119. Bak S, Kahn RA, Nielsen HL, *et al.* Cloning of three A-type cytochromes P450, CYP71E1, CYP98, and CYP99 from *Sorghum bicolor* (L.) Moench by a PCR approach and identification by expression in *Escherichia coli* of CYP71E1 as a multifunctional cytochrome P450 in the biosynthesis of the cyanogenic glucoside dhurrin. *Plant Mol Biol*. 1998;36(3):393-05.
120. Nielsen KA, Tattersall DB, Jones PR, *et al.* Metabolon formation in dhurrin biosynthesis. *Phytochemistry*. 2008;69(1):88-98.
121. Wlodarczyk A, Gnanasekaran T, Nielsen AZ, *et al.* Metabolic engineering of light-driven cytochrome P450 dependent pathways into *Synechocystis* sp. PCC 6803. *Metab Eng*. 2016;33:1-11.
122. Tattersall DB, Bak S, Jones PR, *et al.* Resistance to an herbivore through engineered cyanogenic glucoside synthesis. *Science*. 2001;293(5536):1826-28.
123. Department of Agriculture U, Conservation Service NR. Cassava (*Manihot esculenta* Crantz) plant guide 2003 [homepage on the Internet] Available from: [http://plants.usda.gov/plantguide/pdf/cs\\_maes.pdf](http://plants.usda.gov/plantguide/pdf/cs_maes.pdf).
124. Jorgensen K, Morant AV, Morant M, *et al.* Biosynthesis of the cyanogenic glucosides linamarin and lotaustralin in cassava: isolation, biochemical characterization, and expression pattern of CYP71E7, the oxime-metabolizing cytochrome P450 enzyme. *Plant Physiol*. 2011;155(1):282-92.

125. Jorgensen K, Bak S, Busk PK, *et al.* Cassava plants with a depleted cyanogenic glucoside content in leaves and tubers. Distribution of cyanogenic glucosides, their site of synthesis and transport, and blockage of the biosynthesis by RNA interference technology. *Plant Physiol.* 2005;139(1):363-74.
126. Chen W, Lee M-K, Jefcoate C, *et al.* Fungal cytochrome P450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin. *Genome Biol Evol.* 2014;6(7):1620-34.
127. Ichinose H, Wariishi H. Heterologous expression and mechanistic investigation of a fungal cytochrome P450 (CYP5150A2): involvement of alternative redox partners. *Arch Biochem Biophys.* 2012;518(1):8-15.
128. Shoun H, Takaya N. Cytochromes P450<sub>nor</sub> and P450<sub>foxy</sub> of the fungus *Fusarium oxysporum*. *Int Congr Ser.* 2002;1233:89-97.
129. Durairaj P, Hur JS, Yun H. Versatile biocatalysis of fungal cytochrome P450 monooxygenases. *Microb Cell Fact.* 2016;15(1):125.
130. Lah L, Podobnik B, Novak M, *et al.* The versatility of the fungal cytochrome P450 monooxygenase system is instrumental in xenobiotic detoxification. *Mol Microbiol.* 2011;81(5):1374-89.
131. Julsing MK, Cornelissen S, Bühler B, *et al.* Heme-iron oxygenases: powerful industrial biocatalysts? *Curr Opin Chem Biol.* 2008;12(2):177-86.
132. Murashige T, Skoog F. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiol Plant.* 1962;15(3):473-97.
133. Guengerich FP, Martin MV, Sohl CD, *et al.* Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase. *Nat Protoc.* 2009;4(9):1245-51.
134. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 1970;227(5259):680-85.
135. Li YC, Chiang JY. The expression of a catalytically active cholesterol 7 alpha-hydroxylase cytochrome P450 in *Escherichia coli*. *J Biol Chem.* 1991;266(29):19186-91.
136. Barnes HJ, Arlotto MP, Waterman MR. Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli*. *P Natl Acad Sci USA.* 1991;88(13):5597-01.
137. Barnes HJ. Maximizing expression of eukaryotic cytochrome P450s in *Escherichia coli*. *Methods Enzymol.* 1996;272:3-14.

138. Fisher CW, Caudle DL, Martin-Wixtrom C, *et al.* High-level expression of functional human cytochrome P450 1A2 in *Escherichia coli*. *Faseb J.* 1992;6(2):759-64.
139. Richardson TH, Hsu MH, Kronbach T, *et al.* Purification and characterization of recombinant-expressed cytochrome P450 2C3 from *Escherichia coli*: 2C3 encodes the 6 beta-hydroxylase deficient form of P450 3b. *Arch Biochem Biophys.* 1993;300(1):510-16.
140. Sandhu P, Baba T, Guengerich FP. Expression of modified cytochrome P450 2C10 (2C9) in *Escherichia coli*, purification, and reconstitution of catalytic activity. *Arch Biochem Biophys.* 1993;306(2):443-50.
141. von Wachenfeldt C, Richardson TH, Cosme J, *et al.* Microsomal P450 2C3 is expressed as a soluble dimer in *Escherichia coli* following modification of its N-terminus. *Arch Biochem Biophys.* 1997;339(1):107-14.
142. Rowland P, Blaney FE, Smyth MG, *et al.* Crystal structure of human cytochrome P450 2D6. *J Biol Chem.* 2006;281(11):7614-22.
143. Gay SC, Shah MB, Talakad JC, *et al.* Crystal structure of a cytochrome P450 2B6 genetic variant in complex with the inhibitor 4-(4-Chlorophenyl)imidazole at 2.0-Å resolution. *Mol Pharmacol.* 2010;77(4):529-38.
144. Schoch GA, Yano JK, Wester MR, *et al.* Structure of human microsomal cytochrome P450 2C8. Evidence for a peripheral fatty acid binding site. *J Biol Chem.* 2004;279(10):9497-03.
145. Wester MR, Yano JK, Schoch GA, *et al.* The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-Å resolution. *J Biol Chem.* 2004;279(34):35630-37.
146. Cosme J, Johnson EF. Engineering microsomal cytochrome P450 2C5 to be a soluble, monomeric enzyme. Mutations that alter aggregation, phospholipid dependence of catalysis, and membrane binding. *J Biol Chem.* 2000;275(4):2545-53.
147. Williams PA, Cosme J, Sridhar V, *et al.* Microsomal cytochrome P450 2C5: comparison to microbial P450s and unique features. *J Inorg Biochem.* 2000;81(3):183-90.
148. Strushkevich N, Usanov SA, Plotnikov AN, *et al.* Structural analysis of CYP2R1 in complex with vitamin D3. *J Mol Biol.* 2008;380(1):95-106.
149. Lee DS, Nioche P, Hamberg M, *et al.* Structural insights into the evolutionary paths of oxylipin biosynthetic enzymes. *Nature.* 2008;455(7211):363-68.

150. Kim YH, Kwon T, Yang HJ, *et al.* Gene engineering, purification, crystallization and preliminary X-ray diffraction of cytochrome P450 *p*-coumarate-3-hydroxylase (C3H), the Arabidopsis membrane protein. *Protein Expr Purif.* 2011;79(1):149-55.
151. Chang Z, Wang X, Wei R, *et al.* Functional expression and purification of CYP93C20 a plant membrane-associated cytochrome P450 from *Medicago truncatula*. *Protein Expr Purif.* 2010. DOI: 10.1016/j.pep.2010.11.012
152. Schückel J. Development of a new platform technology for plant Cytochrome P450 fusions. 2012, University of York: PhD thesis.
153. Haudenschild C, Schalk M, Karp F, *et al.* Functional expression of regiospecific cytochrome P450 limonene hydroxylases from mint (*Mentha* spp.) in *Escherichia coli* and *Saccharomyces cerevisiae*. *Arch Biochem Biophys.* 2000;379(1):127-36.
154. Russell DW, Conn EE. The cinnamic acid 4-hydroxylase of pea seedlings. *Arch Biochem Biophys.* 1967;122(1):256-58.
155. Potts JR, Weklych R, Conn EE, *et al.* The 4-hydroxylation of cinnamic acid by sorghum microsomes and the requirement for cytochrome P-450. *J Biol Chem.* 1974;249(16):5019-26.
156. Gabriac B, Werck-Reichhart D, Teutsch H, *et al.* Purification and immunocharacterization of a plant cytochrome P450: the cinnamic acid 4-hydroxylase. *Arch Biochem Biophys.* 1991;288(1):302-09.
157. Mizutani M, Ward E, DiMaio J, *et al.* Molecular cloning and sequencing of a cDNA encoding mung bean cytochrome P450 (P450C4H) possessing cinnamate 4-hydroxylase activity. *Biochem Biophys Res Commun.* 1993;190(3):875-80.
158. Mizutani M, Ohta D, Sato R. Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from Arabidopsis and its expression manner in planta. *Plant Physiol.* 1997;113(3):755-63.
159. Institut de Biologie Moléculaire des Plantes (IBMP) CNRS C. CYP81D11 Co-expression analysis [homepage on the Internet] Available from: [http://www-ibmp.u-strasbg.fr/~CYPedia/CYP81D11/CoExp\\_CYP81D11\\_Pathways.html](http://www-ibmp.u-strasbg.fr/~CYPedia/CYP81D11/CoExp_CYP81D11_Pathways.html).
160. Ekman DR, Lorenz WW, Przybyla AE, *et al.* SAGE analysis of transcriptome responses in Arabidopsis roots exposed to 2,4,6-trinitrotoluene. *Plant Physiol.* 2003;133(3):1397-06.
161. Lorenz A. Bioengineering transgenic plants to detoxify nitroaromatic explosive compounds. 2007, University of York: PhD thesis.

162. Gandia-Herrero F, Lorenz A, Larson T, *et al.* Detoxification of the explosive 2,4,6-trinitrotoluene in *Arabidopsis*: discovery of bifunctional O- and C-glucosyltransferases. *Plant J.* 2008;56(6):963-74.
163. Beynon ER, Symons ZC, Jackson RG, *et al.* The role of oxophytodienoate reductases in the detoxification of the explosive 2,4,6-trinitrotoluene by *Arabidopsis*. *Plant Physiol.* 2009;151(1):253-61.
164. Gunning V, Tzafestas K, Sparrow H, *et al.* *Arabidopsis* Glutathione Transferases U24 and U25 Exhibit a Range of Detoxification Activities with the Environmental Pollutant and Explosive, 2,4,6-Trinitrotoluene. *Plant Physiol.* 2014;165(2):854-65.
165. Bhadra R, Spangord RJ, Wayment DG, *et al.* Characterization of oxidation products of TNT metabolism in aquatic phytoremediation systems of *Myriophyllum aquaticum*. *Environ Sci Technol*; 1999. p. 3354-61.
166. Iwakami S, Endo M, Saika H, *et al.* Cytochrome P450 CYP81A12 and CYP81A21 are associated with resistance to two acetolactate synthase inhibitors in *Echinochloa phyllopogon*. *Plant Physiol.* 2014;165(2):618-29.
167. Pan G, Zhang X, Liu K, *et al.* Map-based cloning of a novel rice cytochrome P450 gene CYP81A6 that confers resistance to two different classes of herbicides. *Plant Mol Biol.* 2006;61(6):933-43.
168. Liu X, Xu X, Li B, *et al.* RNA-Seq transcriptome analysis of maize inbred carrying nicosulfuron-tolerant and nicosulfuron-susceptible alleles. *Int J Mol Sci.* 2015;16(3):5975-89.
169. Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc.* 2007;2(1):31-34.
170. Urban P, Mignotte C, Kazmaier M, *et al.* Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. *J Biol Chem.* 1997;272(31):19176-86.
171. Pompon D, Louerat B, Bronine A, *et al.* Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.* 1996;272:51-64.
172. UniProt: a hub for protein information. *Nucleic Acids Res.* 2015;43(Database issue):D204-12.
173. Nelson D. Cytochrome P450 homepage - P450 STATS 2016 [homepage on the Internet] Available from: <http://drnelson.uthsc.edu/P450.statsfile.html>.

174. Soucek P. Expression of cytochrome P450 2A6 in *Escherichia coli*: purification, spectral and catalytic characterization, and preparation of polyclonal antibodies. *Arch Biochem Biophys*. 1999;370(2):190-00.
175. Bhatt MR, Khatri Y, Rodgers RJ, *et al*. Role of cytochrome b5 in the modulation of the enzymatic activities of cytochrome P450 17alpha-hydroxylase/17,20-lyase (P450 17A1). *J Steroid Biochem Mol Biol*. 2016. doi:10.1016/j.jsbmb.2016.02.033
176. Hatakeyama M, Kitaoka T, Ichinose H. Heterologous expression of fungal cytochromes P450 (CYP5136A1 and CYP5136A3) from the white-rot basidiomycete *Phanerochaete chrysosporium*: Functionalization with cytochrome b5 in *Escherichia coli*. *Enzyme Microb Technol*. 2016;89:7-14.
177. Schenkman JB, Jansson I. The many roles of cytochrome b5. *Pharmacol Ther*. 2003;97(2):139-52.
178. Cabello-Hurtado F, Taton M, Forthoffer N, *et al*. Optimized expression and catalytic properties of a wheat obtusifoliol 14alpha-demethylase (CYP51) expressed in yeast. Complementation of erg11Delta yeast mutants by plant CYP51. *Eur J Biochem*. 1999;262(2):435-46.
179. Schopfer CR, Kochs G, Lottspeich F, *et al*. Molecular characterization and functional expression of dihydroxypterocarpan 6a-hydroxylase, an enzyme specific for pterocarpanoid phytoalexin biosynthesis in soybean (*Glycine max* L.). *FEBS Lett*. 1998;432(3):182-86.
180. Robineau T, Batard Y, Nedelkina S, *et al*. The chemically inducible plant cytochrome P450 CYP76B1 actively metabolizes phenylureas and other xenobiotics. *Plant Physiol*. 1998;118(3):1049-56.
181. Cabello-Hurtado F, Batard Y, Salaun JP, *et al*. Cloning, expression in yeast, and functional characterization of CYP81B1, a plant cytochrome P450 that catalyzes in-chain hydroxylation of fatty acids. *J Biol Chem*. 1998;273(13):7260-67.
182. Helliwell CA, Chandler PM, Poole A, *et al*. The CYP88A cytochrome P450, entkaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc Natl Acad Sci U S A*. 2001;98(4):2065-70.
183. Daly R, Hearn MT. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J Mol Recognit*. 2005;18(2):119-38.
184. Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev*. 2000;24(1):45-66.
185. Macauley-Patrick S, Fazenda ML, McNeil B, *et al*. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*. 2005;22(4):249-70.

186. Andersen MD, Busk PK, Svendsen I, *et al.* 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 recombinant enzymes. *J Biol Chem.* 2000;275(3):1966-75.
187. Katsumata T, Hasegawa A, Fujiwara T, *et al.* Arabidopsis CYP85A2 catalyzes lactonization reactions in the biosynthesis of 2-deoxy-7-oxalactone brassinosteroids. *Biosci Biotechnol Biochem.* 2008;72(8):2110-17.
188. Atkins WM, Sligar SG. Metabolic switching in cytochrome P-450cam: deuterium isotope effects on regiospecificity and the monooxygenase/oxidase ratio. *J Am Chem Soc.* 1987;109(12):3754-60.
189. Murakami H, Yabusaki Y, Sakaki T, *et al.* A genetically engineered P450 monooxygenase: construction of the functional fused enzyme between rat cytochrome P450c and NADPH-cytochrome P450 reductase. *DNA.* 1987;6(3):189-97.
190. Hotze M, Schroder G, Schroder J. Cinnamate 4-hydroxylase from *Catharanthus roseus*, and a strategy for the functional expression of plant cytochrome P450 proteins as translational fusions with P450 reductase in *Escherichia coli*. *FEBS Lett.* 1995;374(3):345-50.
191. Schroder G, Unterbusch E, Kaltenbach M, *et al.* Light-induced cytochrome P450-dependent enzyme in indole alkaloid biosynthesis: tabersonine 16-hydroxylase. *FEBS Lett.* 1999;458(2):97-02.
192. Didierjean L, Gondet L, Perkins R, *et al.* Engineering herbicide metabolism in tobacco and Arabidopsis with CYP76B1, a cytochrome P450 enzyme from Jerusalem artichoke. *Plant Physiol.* 2002;130(1):179-89.
193. Leonard E, Koffas MA. Engineering of artificial plant cytochrome P450 enzymes for synthesis of isoflavones by *Escherichia coli*. *Appl Environ Microbiol.* 2007;73(22):7246-51.
194. Govindaraj S, Poulos TL. Probing the structure of the linker connecting the reductase and heme domains of cytochrome P450BM-3 using site-directed mutagenesis. *Protein Sci.* 1996;5(7):1389-93.
195. Robin A, Roberts GA, Kisch J, *et al.* Engineering and improvement of the efficiency of a chimeric [P450cam-RhFRed reductase domain] enzyme. *Chem Commun.* 2009(18):2478-80.
196. Kulig JK, Spandolf C, Hyde R, *et al.* A P450 fusion library of heme domains from *Rhodococcus jostii* RHA1 and its evaluation for the biotransformation of drug molecules. *Bioorg Med Chem.* 2015;23(17):5603-09.

197. Mizutani M, Ohta D. Two isoforms of NADPH:cytochrome P450 reductase in *Arabidopsis thaliana*. Gene structure, heterologous expression in insect cells, and differential regulation. *Plant Physiol.* 1998;116(1):357-67.
198. Soitamo AJ, Piippo M, Allahverdiyeva Y, *et al.* Light has a specific role in modulating Arabidopsis gene expression at low temperature. *BMC Plant Biol.* 2008;8:13.
199. Sundin L, Vanholme R, Geerinck J, *et al.* Mutation of the inducible *Arabidopsis thaliana* cytochrome P450 reductase 2 alters lignin composition and improves saccharification. *Plant Physiol.* 2014;166(4):1956-71.
200. Varadarajan J, Guilleminot J, Saint-Jore-Dupas C, *et al.* ATR3 encodes a diflavin reductase essential for Arabidopsis embryo development. *New Phytol.* 2010;187(1):67-82.
201. Vincentelli R, Canaan S, Offant J, *et al.* Automated expression and solubility screening of His-tagged proteins in 96-well format. *Anal Biochem.* 2005;346(1):77-84.
202. Kapust RB, Waugh DS. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 1999;8(8):1668-74.
203. Braun P, Hu Y, Shen B, Halleck A, *et al.* Proteome-scale purification of human proteins from bacteria. *Proc Natl Acad Sci USA.* 2002;99(5):2654-59.
204. Dyson MR, Shadbolt SP, Vincent KJ, *et al.* Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. *BMC Biotechnol.* 2004;4:32.
205. Raran-Kurussi S, Waugh DS. The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. *PLoS ONE.* 2012;7(11):e49589.
206. Ray MV, Van Duyne P, Bertelsen AH, *et al.* Production of recombinant salmon calcitonin by in vitro amidation of an *Escherichia coli* produced precursor peptide. *Biotechnology (N Y).* 1993;11(1):64-70.
207. Smith DB. Generating fusions to glutathione S-transferase for protein studies. *Methods Enzymol.* 2000;326:254-70.
208. Swartz JR, Yang J, Voloshin AM, *et al.* Immunogenic protein constructs. Google Patents WO 2008002663 A2. 2013.
209. Nakayama M, Ohara O. A system using convertible vectors for screening soluble recombinant proteins produced in *Escherichia coli* from randomly fragmented cDNAs. *Biochem Biophys Res Commun.* 2003;312(3):825-30.

210. Zhuang R, Zhang Y, Zhang R, *et al.* Purification of GFP fusion proteins with high purity and yield by monoclonal antibody-coupled affinity column chromatography. *Protein Expr Purif.* 2008;59(1):138-43.
211. Schuckel J, Rylott EL, Grogan G, *et al.* A gene-fusion approach to enabling plant cytochromes p450 for biocatalysis. *Chembiochem.* 2012;13(18):2758-63.
212. Chen H, Morgan JA. High throughput screening of heterologous P450 whole cell activity. *Enzyme Microb Technol.* 2006;38(6):760-64.
213. De Mot R, Parret AH. A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes. *Trends Microbiol.* 2002;10(11):502-08.
214. Gustafsson MC, Roitel O, Marshall KR, *et al.* Expression, purification, and characterization of *Bacillus subtilis* cytochromes P450 CYP102A2 and CYP102A3: flavocytochrome homologues of P450 BM3 from *Bacillus megaterium*. *Biochemistry.* 2004;43(18):5474-87.
215. Nakayama N, Takemae A, Shoun H. Cytochrome P450foxy, a catalytically self-sufficient fatty acid hydroxylase of the fungus *Fusarium oxysporum*. *J Biochem.* 1996;119(3):435-40.
216. Winzer T, Kern M, King AJ, *et al.* Plant science. Morphinan biosynthesis in opium poppy requires a P450-oxidoreductase fusion protein. *Science.* 2015;349(6245):309-12.
217. Chak KF, James R. Characterization of the ColE9-J plasmid and analysis of its genetic organization. *J Gen Microbiol.* 1986;132(1):61-70.
218. Su Y, Zou Z, Feng S, *et al.* The acidity of protein fusion partners predominantly determines the efficacy to improve the solubility of the target proteins expressed in *Escherichia coli*. *J Biotechnol.* 2007;129(3):373-82.
219. Wadsater M, Laursen T, Singha A, *et al.* Monitoring shifts in the conformation equilibrium of the membrane protein cytochrome P450 reductase (POR) in nanodiscs. *J Biol Chem.* 2012;287(41):34596-03.
220. Nelson DR. The cytochrome p450 homepage. *Hum Genomics.* 2009;4(1):59-65.
221. Morant M, Bak S, Moller BL, *et al.* Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. *Curr Opin Biotechnol.* 2003;14(2):151-62.
222. Berteau CM, Freije JR, van der Woude H, *et al.* Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua*. *Planta Med.* 2005;71(01):40-47.

223. Teoh KH, Polichuk DR, Reed DW, *et al.* *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Lett.* 2006;580(5):1411-16.
224. Powles SB, Yu Q. Evolution in action: plants resistant to herbicides. *Annu Rev Plant Biol.* 2010;61:317-47.
225. Frear DS. Microsomal N-demethylation, by a cotton leaf oxidase system, of 3-(4'-chlorophenyl)-1, 1-dimethylurea (monuron). *Science.* 1968;162(3854):674-75.
226. Frear DS, Swanson HR, Tanaka FS. N-demethylation of substituted 3-(phenyl)-1-methylureas: Isolation and characterization of a microsomal mixed function oxidase from cotton. *Phytochemistry.* 1969;8(11):2157-69.
227. Mougín C, Cabanne F, Canivenc M-C, *et al.* Hydroxylation and N-demethylation of chlorotoluron by wheat microsomal enzymes. *Plant Sci.* 1990;66(2):195-03.
228. Fonne-Pfister R, Simon A, Salaun J-P, *et al.* Xenobiotic metabolism in higher plants involvement of microsomal cytochrome P-450 in aminopyrine N-demethylation. *Plant Sci.* 1988;55(1):9-20.
229. Batard Y, LeRet M, Schalk M, *et al.* Molecular cloning and functional expression in yeast of CYP76B1, a xenobiotic-inducible 7-ethoxycoumarin O-de-ethylase from *Helianthus tuberosus*. *Plant J.* 1998;14(1):111-20.
230. Abhilash PC, Jamil S, Singh N. Transgenic plants for enhanced biodegradation and phytoremediation of organic xenobiotics. *Biotechnology Adv.* 2009;27(4):474-88.
231. Guengerich FP, Kim DH, Iwasaki M. Role of human cytochrome P-450 IIE1 in the oxidation of many low molecular weight cancer suspects. *Chem Res Toxicol.* 1991;4(2):168-79.
232. Doty SL, Shang TQ, Wilson AM, *et al.* Enhanced metabolism of halogenated hydrocarbons in transgenic plants containing mammalian cytochrome P450 2E1. *Proc Natl Acad Sci USA.* 2000;97(12):6287-91.
233. Doty SL, James CA, Moore AL, *et al.* Enhanced phytoremediation of volatile environmental pollutants with transgenic trees. *Proc Natl Acad Sci USA.* 2007;104(43):16816-21.
234. Zhang Y, Liu J. Transgenic alfalfa plants co-expressing glutathione S-transferase (GST) and human CYP2E1 show enhanced resistance to mixed contaminants of heavy metals and organic pollutants. *J Hazard Mater.* 2011;189(1-2):357-62.

235. Zhang Y, Liu J, Zhou Y, *et al.* Enhanced phytoremediation of mixed heavy metal (mercury)–organic pollutants (trichloroethylene) with transgenic alfalfa co-expressing glutathione S-transferase and human P450 2E1. *J Hazard Mater.* 2013;260:1100-07.
236. Rylott EL, Bruce NC. Plants disarm soil: engineering plants for the phytoremediation of explosives. *Trends Biotechnol.* 2009;27(2):73-81.
237. Jackson RG, Rylott EL, Fournier D, *et al.* Exploring the biochemical properties and remediation applications of the unusual explosive-degrading P450 system XplA/B. *Proc Natl Acad Sci USA.* 2007;104(43):16822-27.
238. Sabbadin F, Jackson R, Haider K, *et al.* The 1.5-A structure of XplA-heme, an unusual cytochrome P450 heme domain that catalyzes reductive biotransformation of royal demolition explosive. *J Biol Chem.* 2009;284(41):28467-75.
239. Hannink N, Rosser SJ, French CE, *et al.* Phytodetoxification of TNT by transgenic plants expressing a bacterial nitroreductase. *Nat Biotechnol.* 2001;19(12):1168-72.
240. Rylott EL, Lorenz A, Bruce NC. Biodegradation and biotransformation of explosives. *Curr Opin Biotechnol.* 2011;22(3):434-40.
241. Rylott EL, Johnston EJ, Bruce NC. Harnessing microbial gene pools to remediate persistent organic pollutants using genetically modified plants--a viable technology? *J Exp Bot.* 2015;66(21):6519-33.
242. Hawari J, Beaudet S, Halasz A, *et al.* Microbial degradation of explosives: biotransformation versus mineralization. *Appl Microbiol Biotechnol.* 2000;54(5):605-18.
243. Fuller ME, Manning JF, Jr. Aerobic gram-positive and gram-negative bacteria exhibit differential sensitivity to and transformation of 2,4,6-trinitrotoluene (TNT). *Curr Microbiol.* 1997;35(2):77-83.
244. Boopathy R, Kulpa CF. Nitroaromatic compounds serve as nitrogen source for *Desulfovibrio* sp. (B strain). *Can J Microbiol.* 1993;39(4):430-33.
245. Bumpus JA, Tatarko M. Biodegradation of 2,4,6-trinitrotoluene by *Phanerochaete chrysosporium*: Identification of initial degradation products and the discovery of a TNT metabolite that inhibits lignin peroxidases. *Curr Microbiol.* 1994;28(3):185-90.
246. Hodgson J, Rho D, Guiot SR, *et al.* Tween 80 enhanced TNT mineralization by *Phanerochaete chrysosporium*. *Can J Microbiol.* 2000;46(2):110-18.
247. Kim HY, Song HG. Transformation and mineralization of 2,4,6-trinitrotoluene by the white rot fungus *Irpex lacteus*. *Appl Microbiol Biotechnol.* 2003;61(2):150-6.

248. Van Aken Bt, Godefroid LM, Peres CM, *et al.* Mineralization of <sup>14</sup>C-U-ring labeled 4-hydroxylamino-2,6-dinitrotoluene by manganese-dependent peroxidase of the white-rot basidiomycete *Phlebia radiata*. *J Biotechnol.* 1999;68(2–3):159-69.
249. Scheibner K, Hofrichter M, Herre A, Michels J, Fritsche W. Screening for fungi intensively mineralizing 2,4,6-trinitrotoluene. *Appl Microbiol Biotechnol.* 1997;47(4):452-57.
250. Bruns-Nagel D, Schmidt TC, Drzyzga O, *et al.* Identification of oxidized TNT metabolites in soil samples of a former ammunition plant. *Environ Sci Pollut R.* 1999;6(1):7-10.
251. Ehltng J, Sauveplane V, Olry A, *et al.* An extensive (co-)expression analysis tool for the cytochrome P450 superfamily in *Arabidopsis thaliana*. *BMC Plant Biol.* 2008;8:47.
252. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods.* 2012;9(7):671-75.
253. Katzen F. Gateway((R)) recombinational cloning: a biological operating system. *Expert Opin Drug Discov.* 2007;2(4):571-89.
254. Voinnet O, Rivas S, Mestre P, *et al.* An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* 2003;33(5):949-56.
255. Bernhardt R, Urlacher VB. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl Microbiol Biotechnol.* 2014;98(14):6185-03.
256. Tzafestas K. Investigating the role of glutathione transferases in the phytodetoxification of explosives. 2016, University of York: PhD thesis.
257. Mezzari MP, Walters K, Jelínková M, *et al.* Gene expression and microscopic analysis of *Arabidopsis* exposed to chloroacetanilide herbicides and explosive compounds. A phytoremediation approach. *Plant Physiol.* 2005;138(2):858-69.
258. Rylott EL, Gunning V, Tzafestas K, *et al.* Phytodetoxification of the environmental pollutant and explosive 2,4,6-trinitrotoluene. *Plant Signal Behav.* 2015;10(1):e977714.
259. Shinkai Y, Nishihara Y, Amamiya M, *et al.* NADPH-cytochrome P450 reductase-mediated denitration reaction of 2,4,6-trinitrotoluene to yield nitrite in mammals. *Free Radic Biol Med.* 2016;91:178-87.
260. Esteve-Nunez A, Caballero A, Ramos JL. Biological degradation of 2,4,6-trinitrotoluene. *Microbiol Mol Biol Rev.* 2001;65(3):335-52.

261. Parrish FW. Fungal transformation of 2,4-dinitrotoluene and 2,4,6-trinitrotoluene. *Appl Environ Microbiol.* 1977;34(2):232-33.
262. Sagi-Ben Moshe S, Ronen Z, Dahan O, *et al.* Sequential biodegradation of TNT, RDX and HMX in a mixture. *Environ Pollut.* 2009;157(8-9):2231-38.
263. Baerson SR, Sanchez-Moreiras A, Pedrol-Bonjoch N, *et al.* Detoxification and transcriptome response in Arabidopsis seedlings exposed to the allelochemical benzoxazolin-2(3H)-one. *J Biol Chem.* 2005;280(23):21867-81.
264. Bruce TJ, Matthes MC, Chamberlain K, *et al.* Cis-Jasmone induces Arabidopsis genes that affect the chemical ecology of multitrophic interactions with aphids and their parasitoids. *Proc Natl Acad Sci USA.* 2008;105(12):4553-58.
265. Matthes MC, Bruce TJ, Ton J, *et al.* The transcriptome of cis-jasmone-induced resistance in Arabidopsis thaliana and its role in indirect defence. *Planta.* 2010;232(5):1163-80.
266. Matthes M, Bruce T, Chamberlain K, *et al.* Emerging roles in plant defense for cis-jasmone-induced cytochrome P450 CYP81D11. *Plant Signal Behav.* 2011;6(4):563-65.
267. Mueller S, Hilbert B, Dueckershoff K, *et al.* General detoxification and stress responses are mediated by oxidized lipids through TGA transcription factors in Arabidopsis. *Plant Cell.* 2008;20(3):768-85.
268. Walper E, Weiste C, Mueller MJ, *et al.* Screen identifying Arabidopsis transcription factors involved in the response to 9-lipoxygenase-derived oxylipins. *PLoS One.* 2016;11(4):e0153216.
269. Bernhardt R. Cytochromes P450 as versatile biocatalysts. *J Biotechnol.* 2006;124(1):128-45.
270. Chang MC, Eachus RA, Trieu W, *et al.* Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat Chem Biol.* 2007;3(5):274-77.
271. Bollag DM, McQueney PA, Zhu J, *et al.* Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action. *Cancer Res.* 1995;55(11):2325-33.
272. Werck-Reichhart D, Hehn A, Didierjean L. Cytochromes P450 for engineering herbicide tolerance. *Trends Plant Sci.* 2000;5(3):116-23.