Development of
a new platform technology for plant
Cytochrome P450 fusions

Julia Schückel
PhD Thesis

University of York
Department of Biology

January 2012
Abstract

To date more than 15000 Cytochromes P450 have been identified and named so far, with one third belonging to the plant kingdom. This is a key biochemical resource, providing a wealth of biocatalysts covering a diverse range of chemistries. Characterisation, however, has been greatly hindered by the poor solubility of many P450s, a result of the membrane anchoring region common to all plant P450s. Fusions of plant P450 heme domains to an appropriate reductase without the hydrophobic membrane anchor could provide the basis for developing robust, soluble plant enzyme systems for substrate screens to discover novel activities that are also of benefit to industry.

In this project, the two predominantly expressed Arabidopsis reductases ATR1 and ATR2 have been cloned, without the membrane anchor, and expressed in *Escherichia coli*. These two truncated enzymes have been purified and assessed for activity with ATR2 found to be more active than ATR1. ATR2 was chosen for engineering into a novel plant P450 reductase vector platform for high throughput applications, whereby the P450s can be easily and quickly swapped using ligation independent cloning techniques. Four different plant P450s (CYP93C1, CYP73A5, CYP82E4 and CYP81D8) were selected to validate this technology, and activity for the fusions of CYP93C1 (Isoflavone synthase I from *Glycine max*) and CYP73A5 (cinnamate-4-hydroxylase from Arabidopsis) with ATR2 have been shown. The presence of CYP73A5 fused to ATR2 was verified through purification and further studies showed that it has to be membrane associated for activity.

Additionally, CYP93C1 and CYP73A5 were also fused with the bacterial RhF reductase from *Rhodococcus* sp. and expressed in *E. coli* and compared to the plant P450 – plant reductase fusion protein. These novel plant-bacterial fusion P450 systems are the first example of active plant P450s fused to a reductase from a bacterium.

This platform technology will provide the possibility for characterisation studies of eukaryotic P450s with unknown function and the discovery of new activities.
Abstract .............................................................................................................. 2
List of Tables ........................................................................................................ 8
List of Figures ......................................................................................................... 9
Acknowledgements .............................................................................................. 14
Author’s Declaration ............................................................................................ 15

1 Introduction ....................................................................................................... 16
1.1 Biotransformations in Industry .................................................................. 16
1.2 Cytochromes P450 ..................................................................................... 17
  1.2.1 History and Nomenclature of Cytochromes P450 .......................... 17
  1.2.2 Structure and Function of Cytochromes P450 .............................. 22
  1.2.3 Reaction Mechanism ...................................................................... 25
  1.2.4 Cytochromes P450 reactions .......................................................... 27
  1.2.5 Interactions between Cytochromes P450 and reductases ............ 28
  1.2.6 Cytochrome P450 Fusion Systems ................................................. 29
1.3 Applications of Cytochromes P450 in biocatalysis ......................... 33
  1.3.1 Hydroxylation .............................................................................. 33
  1.3.2 Oxidation ...................................................................................... 35
1.4 Plant Cytochromes P450 ........................................................................ 37
  1.4.1 Cytochromes P450s in Arabidopsis thaliana .............................. 37
  1.4.2 Fusions of plant Cytochromes P450 ............................................ 43
1.5 Recombinant Expression Systems .......................................................... 47
  1.5.1 Expression of P450s in Escherichia coli .................................... 48
  1.5.2 Expression in Yeast .................................................................... 48
1.6 Aim of the project ...................................................................................... 50

2 Materials and Methods ............................................................................... 51
2.1 Reagents and consumables ..................................................................... 51
2.2 Media, strains and plasmids .................................................................... 52
  2.2.1 Bacterial media ............................................................................ 52
  2.2.2 Bacterial strains .......................................................................... 53
  2.2.3 Plasmids for gene cloning and enzyme expression .................... 55
2.3 Preparation of chemically competent *Escherichia coli* cells .......... 58
2.4 Plant media, growth conditions and strains .................................. 59
2.5 Plasmid DNA preparation ...................................................................... 60
2.6 Purification of DNA fragments ............................................................... 60
2.7 Detection of the mRNA transcript ......................................................... 60
  2.7.1 RNA extraction from Escherichia coli ....................................... 60
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7.2 RNA extraction from plant tissue</td>
<td>61</td>
</tr>
<tr>
<td>2.7.3 Reverse transcription from plant mRNA to cDNA</td>
<td>62</td>
</tr>
<tr>
<td><strong>2.8 Traditional cloning method</strong></td>
<td>62</td>
</tr>
<tr>
<td>2.8.1 Designing primers for truncated P450 and reductase genes</td>
<td>62</td>
</tr>
<tr>
<td>2.8.2 Polymerase chain reaction (PCR)</td>
<td>62</td>
</tr>
<tr>
<td>2.8.3 Preparation of PCR products for cloning</td>
<td>63</td>
</tr>
<tr>
<td>2.8.4 Transformation of plasmid DNA into chemically competent Escherichia coli cells</td>
<td>63</td>
</tr>
<tr>
<td>2.8.5 Screening transformants for recombinant genes by PCR</td>
<td>64</td>
</tr>
<tr>
<td>2.8.6 Restriction endonuclease digest of DNA</td>
<td>64</td>
</tr>
<tr>
<td>2.8.7 Dephosphorylation of DNA 5’ end of the vector</td>
<td>64</td>
</tr>
<tr>
<td>2.8.8 DNA ligation reaction</td>
<td>65</td>
</tr>
<tr>
<td>2.8.9 DNA sequencing and analysis</td>
<td>65</td>
</tr>
<tr>
<td><strong>2.9 Ligation independent cloning method</strong></td>
<td>66</td>
</tr>
<tr>
<td><strong>2.10 Protein expression and purification</strong></td>
<td>67</td>
</tr>
<tr>
<td>2.10.1 Expression in Escherichia coli</td>
<td>67</td>
</tr>
<tr>
<td>2.10.2 Solubilisation buffers</td>
<td>68</td>
</tr>
<tr>
<td>2.10.3 Cell lysis by sonication</td>
<td>70</td>
</tr>
<tr>
<td>2.10.4 Protein purification in a batch process</td>
<td>71</td>
</tr>
<tr>
<td>2.10.5 Protein purification in a continuous process</td>
<td>71</td>
</tr>
<tr>
<td><strong>2.11 Agarose gel electrophoresis</strong></td>
<td>72</td>
</tr>
<tr>
<td><strong>2.12 Protein detection</strong></td>
<td>72</td>
</tr>
<tr>
<td><strong>2.13 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)</strong></td>
<td>72</td>
</tr>
<tr>
<td>2.13.1 SDS PAGE with the Bio-Rad system</td>
<td>72</td>
</tr>
<tr>
<td>2.13.2 SDS PAGE with the RunBlue system from Expedeon</td>
<td>73</td>
</tr>
<tr>
<td><strong>2.14 Western blot analysis</strong></td>
<td>73</td>
</tr>
<tr>
<td>2.14.1 Protein detection with anti-poly Histidine peroxidase conjugate</td>
<td>73</td>
</tr>
<tr>
<td>2.14.2 Protein detection with alkaline phosphatase conjugate</td>
<td>74</td>
</tr>
<tr>
<td><strong>2.15 Protein characterisation</strong></td>
<td>75</td>
</tr>
<tr>
<td>2.15.1 Characterisation of the P450 reductases</td>
<td>75</td>
</tr>
<tr>
<td>2.15.1.1 Spectrophotometric characterisation</td>
<td>75</td>
</tr>
<tr>
<td>2.15.1.2 Activity assay with cytochrome c</td>
<td>75</td>
</tr>
<tr>
<td>2.15.1.3 Temperature and pH optima of ATR1tr and ATR2tr</td>
<td>76</td>
</tr>
<tr>
<td>2.15.1.4 Stability test of ATR2tr in different buffers</td>
<td>77</td>
</tr>
<tr>
<td>2.15.1.5 Kinetic studies of the reductases</td>
<td>77</td>
</tr>
<tr>
<td>2.15.2 Protein identification by MALDI-MS analysis</td>
<td>78</td>
</tr>
<tr>
<td>2.15.3 P450 enzyme activity assays</td>
<td>78</td>
</tr>
<tr>
<td>2.15.3.1 Activity assay for CYP71D15 with limonene</td>
<td>78</td>
</tr>
</tbody>
</table>
## Table of Contents

2.15.3.2 Activity assay for CYP81D8 and CYP81D11 with methyl-tolyl-sulphide ................................................................. 79
2.15.3.3 Isoflavone synthase activity assay ......................................................... 80
2.15.3.4 Cinnamate-4-hydroxylase (CYP73A5) activity assay ............... 81
2.15.3.5 Activity assay for CYP82E4 ................................................................. 82
2.15.3.6 Activity assay for CYP81D8 with TNT and aminodinitrotoluene 83
2.15.3.7 Activity assay for CYP81D8 with 7-ethoxycoumarin ............... 84

3 Chapter: Expression of Cytochromes P450 .................. 86
3.1 Introduction ............................................................................................................. 86
3.2 Objectives ............................................................................................................. 90
3.3 Materials and Methods ..................................................................................... 90
  3.3.1 Analysis of CYP73D15 PM2-2 construct .......................................................... 90
  3.3.2 Materials and Methods for P450 expression in Escherichia coli 92
  3.3.2.1 Isolation and cloning of the plant P450s .................................................. 92
  3.3.3 Expression of PM2-2 in Escherichia coli .......................................................... 92
  3.3.4 Materials and Methods for P450 expression in yeast ................................. 93
  3.3.4.1 Media ............................................................................................................... 93
  3.3.4.2 Transformation and expression in yeast ................................................... 94
  3.3.4.3 Yeast microsome preparation ........................................................................ 95
3.4 Results ................................................................................................................. 95
  3.4.1 P450 amino acid sequence analysis ............................................................... 95
  3.4.2 Analysis of CYP71D15 PM2-2 ........................................................................ 99
  3.4.3 P450 cloning into the LIC-vector .............................................................. 100
  3.4.4 P450 expression in Escherichia coli ............................................................. 101
  3.4.5 Activity assay for 81D8tr, 81D11tr and CYP71D15 PM2-2 .... 105
  3.4.6 P450 expression in yeast ................................................................................ 106
3.5 Discussion ........................................................................................................... 108
  3.5.1 P450s expressed in Escherichia coli ......................................................... 108
  3.5.2 P450s expressed in Saccharomyces cerevisiae .................................. 110

4 Expression, purification and characterisation of truncated, soluble Arabidopsis Cytochrome P450 reductases .............................................................. 111
4.1 Introduction ........................................................................................................... 111
4.2 Objectives ........................................................................................................... 116
4.3 Results ................................................................................................................. 116
  4.3.1 Primer design for the N-terminal membrane anchor truncated ATR1 and ATR2 .......................................................... 116
# Table of Contents

4.3.2 Cloning of ATR1tr and ATR2tr ......................................................... 118
4.3.3 Expression of ATR1tr and ATR2tr in Escherichia coli ............... 118
4.3.4 Solubilisation of overexpressed ATR1tr ..................................... 120
4.3.5 Purification of ATR1tr and ATR2tr in solubilisation buffer 18A ................................................................. 121
4.3.6 Characterisation of Arabidopsis P450 reductase ATR2tr .......... 123
  4.3.6.1 Activity assay with cytochrome c ........................................ 123
  4.3.6.2 Temperature and pH optima of ATR1tr and ATR2tr ............. 126
  4.3.6.3 Kinetic studies of the reductases ATR1tr and ATR2tr .......... 128
  4.3.6.4 Spectrophotometric characterisation of ATR2tr .................. 131
  4.3.6.5 Activity test of purified ATR2tr in buffer A and B .......... 132
  4.3.6.6 Stability test of ATR2tr in different buffers ..................... 133

4.4 Discussion ................................................................................. 134

5 Development of the ATR2tr-LIC platform ......................... 137
  5.1 Introduction ........................................................................ 137
  5.2 Objectives ........................................................................... 140
  5.3 Methods .................................................................................. 141
    5.3.1.1 Developing the platform technology of plant P450-ATR2tr fusions ................................................................. 141
    5.3.2 Cloning of P450s into lamATR2tr- and licATR2tr-vector ..... 142
      5.3.2.1 Primer design ............................................................. 142
      5.3.2.2 Cloning strategy ......................................................... 144
    5.3.3 Expression and activity assays of the novel plant fusions .... 144

5.4 Results ....................................................................................... 145
  5.4.1 Developing the platform technology for plant P450-ATR2tr fusions ................................................................. 145
    5.4.1.1 Designing the linker region between P450 and reductase .... 145
    5.4.2 Cloning of ATR2tr into the LIC-vector ............................. 146
    5.4.3 Cloning of P450 inserts into lamATR2tr and licATR2tr ....... 147
  5.4.4 Expression and activity assays of the novel plant fusions .... 147
    5.4.4.1 IFS-ATR2tr fusions ...................................................... 147
    5.4.4.2 CYP73A5tr-ATR2tr fusions .......................................... 152
    5.4.4.3 CYP82E4tr-ATR2tr fusions .......................................... 157
    5.4.4.4 CYP81D8tr-ATR2tr fusions .......................................... 160

5.5 Discussion ................................................................................. 162

6 Comparison of plant Cytochromes P450 fused to different reductases ......................................................... 167
  6.1 Introduction ........................................................................ 167
# Table of Contents

6.2 Objectives ........................................................................................................... 167
6.3 Methods .............................................................................................................. 168
  6.3.1 Expression of IFS-fusions ........................................................................... 168
    6.3.1.1 Expression conditions of IFS-CPR .................................................. 168
    6.3.1.2 Optimised conditions for the expression of the IFS-RhF fusion ........ 169
6.4 Results ............................................................................................................... 170
  6.4.1 IFS fusions .................................................................................................... 170
    6.4.1.1 Expression of IFS-CPR ................................................................... 170
    6.4.1.2 Expression of IFS-RhF ................................................................... 172
    6.4.1.3 Comparison of the IFS fusions in their growing cell cultures ....... 176
    6.4.1.4 Detection of IFS fusion enzymes ..................................................... 177
    6.4.1.5 Purification of IFS-RhF ................................................................. 180
  6.4.2 CYP73A5tr fusions ....................................................................................... 180
6.5 Discussion .......................................................................................................... 183
7 Final Discussion ..................................................................................................... 188
Appendix A .............................................................................................................. 194
Appendix B .............................................................................................................. 203
Abbreviations .......................................................................................................... 214
References ................................................................................................................ 218
List of Tables

Table 1.1: Arabidopsis P450s with known function .................................................. 39
Table 1.2: Plant fusion P450s heterologously expressed in *E. coli* ...................... 45
Table 2.1: List of *E. coli* strains used for cloning and recombinant expression in this project ........................................................................................................ 54
Table 2.2: Plasmids for gene cloning and enzyme expression .............................. 55
Table 2.3: List of primers used for sequencing ....................................................... 65
Table 2.4: Protocol for T4 polymerase treatment ................................................. 67
Table 2.5: List of solubilisations buffer adapted from Lindwall et al. 2000^347^ .... 69
Table 2.6: Buffers used for analyzing the pH optima of ATR1tr and ATR2tr ....... 76
Table 2.7: HPLC program for the separation of methyl-tolyl-sulphide and its oxidised derivates ................................................................. 80
Table 2.8: HPLC program for the separation cinnamic acid and coumaric acid at 1 ml/min ........................................................................................................... 82
Table 3.1: *Saccharomyces cerevisiae* strains used in this work .......................... 86
Table 3.2: P450s upregulated after TNT treatment from microarray data^363^..... 87
Table 3.3: Primer used for sequence analysis of CYP71D15 PM2-2 ................. 91
Table 3.4: Primer used for PCR amplification of the truncated versions of CYP81D8 and CYP81D11 .......................................................... 92
Table 3.5: Scheme of the test expression of the Arabidopsis P450s in the Lic- vector using different *E. coli* strains and media ........................................ 102
Table 4.1: Kinetic data for ATR1tr and ATR2tr (calculated by the program GraFit) .................................................................................................................. 129
Table 5.1: Primer sequences for cloning ATR2tr into the LIC-vector ............ 141
Table 5.2: List of P450s used for the creation of artificial plant fusions with the reductase ATR2tr .............................................................. 143
Table 5.3: Primer sequences for cloning P450 inserts into ATR2tr-LIC-vector 144
Table 5.4: List of chemical compounds considered to use as internal standard optimising the extraction method for the two flavonoids naringenin and genistein .............................................................................................................. 148
Table 7.1: List of P450-reductase fusion constructs created in this project and their tested substrates ................................................................. 190
List of Figures

Figure 1.1: Original figure of the carbon monoxide difference spectra present in rat liver microsomes (from Klingenberg 1958) ................................................................. 18
Figure 1.2: Scientific articles published on National Center for Biotechnology Information (23th January 2012) ................................................................. 19
Figure 1.3: Schematic distribution of cytochromes P450 among the different kingdoms .............................................................................................................. 19
Figure 1.4: Nomenclature of P450s .................................................................................. 20
Figure 1.5: Schematic organisation of A the plant cytochrome and B the bacteria P450 systems ........................................................................................................ 21
Figure 1.6: Ribbon diagram of Arabidopsis allene oxide synthase (CYP74A1) crystal structure ................................................................................................................. 22
Figure 1.7: Topographic map showing the secondary elements represented for the CYP102 and commonly found among P450 enzymes ........................................ 23
Figure 1.8: Ribbon diagram catalytic P450 centre with the 5-coordinated cysteine ................................................................. 24
Figure 1.9: Reaction cycle of a P450 hydroxylation ........................................................................ 26
Figure 1.10: Variety of reaction types catalysed by cytochromes P450 .......................... 28
Figure 1.11: Schematic arrangement of natural P450-reductase fusion systems ........................................ 31
Figure 1.12: O-dealkylation of 7-ethoxycoumarin to 7-hydroxycoumarin by P450-RhF ...................................................................................................................... 32
Figure 1.13: Hydroxylation of camphor by P450cam (Pseudomonas putida) to 5-exo-hydroxycamphor or P450camr (Rhodococcus sp.) to 6-endo-hydroxycamphor ................................................................................................. 33
Figure 1.14: Hydroxylation of progesterone to deoxycorticosterone by CYP21A2 ...................................................................................................................... 34
Figure 1.15: Hydroxylation of 6-oxo-campestanol to cathasterone by Arabidopsis CYP90B1 ...................................................................................................................... 35
Figure 1.16: P450 dependent heteroatom-oxidations of tertiary amines to N-oxides and thioether to sulfoxides or sulfones ........................................................................... 35
Figure 1.17: P450 catalysed oxidation of 3-methylpyridine to pyridine-3-carboxylic acid ...................................................................................................................... 36
Figure 1.18: P450 catalysed epoxidation of an alkene ........................................................................ 36
Figure 1.19: Phylogenetic tree of all cytochromes P450 from Arabidopsis 38
Figure 1.20: 14α-demethylation of obtusifoliol by plant CYP51 .................................... 43
Figure 1.21: Conversion of the flavanones liquiritigenin and naringenin to the isoflavones daidzein and genistein, respectively, by the P450 isoflavone synthase I ...................................................................................................................... 46
Figure 1.22: Reaction mechanism of the conversion of naringenin to genistein by the P450 isoflavone synthase I ...................................................................................................................... 47
Figure 2.1: Vector map for pTrcHis2/lacZ ................................................................................ 56
Figure 2.2: Vector map of LIC-vector ................................................................................ 57
Figure 2.3: pYeDP60 shuttle vector map with an origin of replication for *E. coli* and yeast ................................................................. 58
Figure 2.4: Schematic procedure for ligation independent cloning ................... 66
Figure 2.5: LIC-vector after T4/dTTPP treatment ..................................... 67
Figure 2.6: *E. coli* Rosetta 2 (DE3) cells A before and B after sonication ....... 70
Figure 2.7: GC-MS program for analysing limonene and its hydroxylated products ................................................................. 79
Figure 2.8: HPLC chromatogram for the separation of nicotine and nornicotine 83
Figure 3.1: Chemical structure of 2,4,6-trinitrotoluene (TNT) .......................... 87
Figure 3.2: Occurence of CYP81D8 and CYP81D11 in the development of *Arabidopsis* ................................................................. 88
Figure 3.3: Regiospecific hydroxylation of (-)-4S-limonene to (-)-trans-isopiperitenol within the biosynthesis of (-)-menthol .......................... 89
Figure 3.4: N-terminal modification of CYP71D15 creating construct PM2-2 367 90
Figure 3.5: Vector map for pCWori+ ......................................................... 91
Figure 3.6: Analysis of the CYP81D8 amino acid sequence .......................... 96
Figure 3.7: Analysis of the CYP81D11 amino acid sequence ........................ 97
Figure 3.8: Analysis of the CYP71D15 amino acid sequence ........................ 98
Figure 3.9: Analysis of CYP71D15 PM2-2 in pCWori+ vector after digestion .. 99
Figure 3.10: Analysis of PCR products for native CYP81D8 (8nat), N-truncated CYP81D8 (8tr), native CYP81D11 (11nat) and N-truncated CYP81D11 (11tr) ................................................................. 100
Figure 3.11: Analysis of crude extract of JM109 cells after expression of CYP71D15 PM2-2 (57 kDa) ................................................................. 101
Figure 3.12: SDS PAGE analysis of crude extract from *E. coli* Rosetta 2 cells expressing 81D8tr (54 kDa) in LB medium with and without induction ... 102
Figure 3.13: SDS PAGE analysis of crude extract from *E. coli* Rosetta 2 cells expressing 81D11tr (53 kDa) in LB medium with and without induction . 103
Figure 3.14: SDS PAGE analysis of protein fractions containing 81D8tr and 81D11tr after purification on Ni-resin material ................................. 103
Figure 3.15: SDS PAGE analysis of protein fractions containing 81D8tr and 81D11tr after purification on Ni-resin material ................................. 104
Figure 3.16: SDS PAGE analysis of protein fractions containing CYP71D15 PM2-2 and empty vector control after purification on Ni-resin material ......................... 104
Figure 3.17: Analysis of plant P450s after purification on Ni-resin material ....... 105
Figure 3.18: HPLC chromatogram of the methyl-tolyl-sulphide assay ................ 106
Figure 3.19: CO difference spectrum of the microsomes for XplA heme from *Rhodococcus* sp. and CYP71D15 from peppermint ........................................ 107
Figure 3.20: CO difference spectrum of the microsomes and the supernatant after the ultracentrifugation for CYP81A11 from *Arabidopsis* and CYP71D15 from peppermint .................................................. 107
Figure 3.21: SDS PAGE analysis of yeast microsomes containing CYP81D11(57 kDa) and CYP71D15 PM2-2 (57 kDa) .............................................. 108
LIST OF FIGURES

Figure 4.1: Phylogenetic tree of eukaryotic cytochrome P450 reductases ........ 113
Figure 4.2: Amino acid sequence alignment between the CPR from C. roseus and
the two Arabidopsis reductases ATR1 and ATR2 modified from Mizutani
and Otha 1998391 ........................................................................................................ 114
Figure 4.3: Expression levels of the three Arabidopsis cytochrome P450
reductases ATR1, ATR2 and ATR3 during the plant life cycle ................. 115
Figure 4.4: Amino acid sequence analysis of ATR1 N-terminus. ............... 117
Figure 4.5: Cell growth (OD600) after induction of ATR1tr in E. coli Rosetta 2
grown in LB medium .......................................................... 118
Figure 4.6: Analysis of crude extract by SDS PAGE ................................... 119
Figure 4.7: SDS PAGE analysis of ATR1tr (72kDa) during Ni-affinity
purification .................................................................................... 120
Figure 4.8: Protein concentration of the soluble total protein after overexpression
in E. coli Rosetta 2 (DE3) using different solubilisation buffers 0-30...... 120
Figure 4.9: SDS PAGE analysis of the total soluble protein using solubilisation
buffer 0-30. Red box show the overexpressed ATR1tr (72 kDa) .............. 121
Figure 4.10: Analysis of ATR1tr after sonication in different solubilisation buffers
........................................................................................................ 121
Figure 4.11: SDS PAGE Analysis of the purification of A) ATR1tr (72 kDa), B)
ATR2tr (71 kDa) and C) negative control (empty vector expressed in E. coli
Rosetta 2 (DE3)) ........................................................................... 122
Figure 4.12: ATR1tr and ATR2tr activity dependence on the cofactor NADH and
NADPH........................................................................................................ 123
Figure 4.13: Cytochrome c conversion by different reductases in presence of the
cofactor NADPH .................................................................................. 124
Figure 4.14: Cytochrome c conversion by different reductases in presence of the
cofactor NADPH .................................................................................. 125
Figure 4.15: Temperature optimum for ATR1tr (1 mg/ml) and ATR2tr
(0.01 mg/ml)................................................................................... 126
Figure 4.16: Activity of 100 μg/ml ATR1tr and 1 μg/ml ATR2tr in 300 mM
Britton Robinson buffer (1:1:1 mixture of boric acid, phosphoric acid and
acetic acid) at different pH values .................................................... 127
Figure 4.17: Activity of 1 μg/ml ATR2tr in diverse buffers covering different pH
values ............................................................................................... 127
Figure 4.18: Stability of the cofactor NADPH at different pH values ........... 128
Figure 4.19: Michaelis-Menten blots by GraFit for A) ATR1tr by Michaelis-
Menten equation and B) ATR1tr by equation for substrate inhibition and C)
ATR2tr by Michaelis-Menten equation ........................................... 129
Figure 4.20: Cornish-Bowden blot for A) ATR1tr and B) ATR2tr ............. 130
Figure 4.21: UV/Vis spectra of purified recombinant ATR2tr and negative control
........................................................................................................ 131
Figure 4.22: UV/Vis spectra of 10 μM FMN and 10 μM FAD in 50 mM
potassium phosphate buffer pH 7.5 ............................................. 131
Figure 4.23: Michaelis-Menten-diagram of purified ATR2tr .................. 132
Figure 4.24: Stability of ATR2tr over time in A) buffer A (50 mM potassium phosphate buffer pH 7.5) and B) buffer B (30 mM potassium phosphate pH 7.8, 20 % glycerol, 0.1 mM EDTA, 2.0 μM FMN) .................................................. 133
Figure 5.1: Phenylpropanoid pathway showing the key role of the two P450s IFS (CYP93C) and C4H (CYP73A5) used in this project ................................................................. 138
Figure 5.2: Arrangement of IFS-CP fusion construct ................................................................. 139
Figure 5.3: N-demethylation of nicotine to nornicotine by CYP82E4 from Nicotiana tabacum .......................................................................................................................... 139
Figure 5.4: Schematic construction of the P450-ATR2tr fusion containing the lam-linker ................................................................................................................................. 145
Figure 5.5: Analysis of positive transformants for the presence of the lam-ATR2tr in the LIC-vector .................................................................................................................. 146
Figure 5.6: Analysis of positive transformants for the presence of the P450 73A5tr .......................................................... ................................................................. 147
Figure 5.7: HPLC analysis of the flavonoids biochanine A, genistein and naringenin ........................................................................................................................................ 148
Figure 5.8: HPLC analysis of the flavonoids scopoletin, naringenin and genistein ........................................................................................................................................ 149
Figure 5.9: Activity of Glycine max IFS fused to Arabidopsis ATR2tr ................................. 150
Figure 5.10: A) SDS-PAGE and B) western blot analysis (anti ATR2-antibodies) of Glycine max IFS fused to Arabidopsis ATR2tr ................................. 151
Figure 5.11: Activity of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr in different media .......................................................................................................................... 152
Figure 5.12: Activity of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr in M9 medium .......................................................................................................................... 153
Figure 5.13: Conversion of cinnamic acid to coumaric acid by artificial plant fusions containing Arabidopsis P450 73A5tr and Arabidopsis ATR2tr in a resting cell assay ................................................................................... 154
Figure 5.14: SDS-PAGE and western blot analysis (anti His-antibody) of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr ................................................................. 155
Figure 5.15: Cinnamate-4-hydroxylase activity in different fractions of the purification of 73A3tr-lamATR2tr ................................................................................................. 155
Figure 5.16: Cinnamate-4-hydroxylase activity of 73A3tr-lamATR2tr in different fractions after expression in E. coli ................................................................................................. 156
Figure 5.17: SDS-PAGE and western blot analysis of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr as soluble and insoluble protein .................................................................... 156
Figure 5.18: Nicotine removal in a resting cell assay ................................................................ 157
Figure 5.19: HPLC analysis (reversed-phase column) of the crude cell extract for nicotine and nornicotine ........................................................................................................ 158
Figure 5.20: HPLC analysis (chiral column) of the crude cell extract for nicotine and nornicotine ........................................................................................................ 159
Figure 5.21: SDS-PAGE and western blot analysis of Nicotiana tabacum P450 82E4tr fused to Arabidopsis ATR2tr ................................................................................................. 159
Figure 5.22: TNT removal in a resting cell assay over time .................................................. 160
Figure 5.23: Resting cell assay using 7-ethoxycoumarin as substrate ................................. 161
Figure 5.24: SDS-PAGE and western blot analysis of Arabidopsis P450 81D8tr fused to Arabidopsis ATR2tr ................................................................. 161
Figure 6.1: Flow scheme of the protocol for the expression of IFS-CPR construct according to Leonard and Koffas306 ............................................. 169
Figure 6.2: Observation of the cell growth of E. coli JM109 and Rosetta 2 during the expression of IFS-CPR .......................................................... 170
Figure 6.3: Activity of Glycine max IFS fused to Catharanthus roseus CPR when expressed with E. coli JM109 and Rosetta 2 cells ......................... 171
Figure 6.4: Activity of Glycine max IFS fused to Rhodococcus sp. RhF reductase .......................................................................................... 172
Figure 6.5: Conversion of naringenin to genistein by IFS-RhF and IFS-CPR in a growing cell assay at 20 °C .......................................................... 173
Figure 6.6: Conversion of naringenin to genistein by IFS-RhF and IFS-CPR in a growing cell assay at 15 °C .......................................................... 174
Figure 6.7: Observation of the optical density during the expression of IFS-CPR and IFS-RhF at 20 °C in M9 medium ........................................ 175
Figure 6.8: Conversion of naringenin to genistein in a resting cell assay ......... 175
Figure 6.9: Production of genistein by IFS fusions in a growing cell assay ...... 176
Figure 6.10: SDS-PAGE and western blot analysis of Glycine max IFS fused to Catharanthus roseus CPR and to Rhodococcus sp. RhF reductase in the soluble and insoluble fraction ........................................ 177
Figure 6.11: SDS-PAGE and western blot analysis of Glycine max IFS fused to the RhF reductase from Rhodococcus sp. in total, soluble and insoluble protein fraction .................................................. 178
Figure 6.12: SDS-PAGE and western blot analysis of Glycine max IFS fused to Arabidopsis reductase ATR2tr in total, soluble and insoluble protein fraction .................................................. 178
Figure 6.13: Reverse transcription PCR for IFS-RhF .................................................. 179
Figure 6.14: SDS-PAGE and western blot analysis of Glycine max IFS fused to Rhodococcus sp. RhF reductase after purification with Ni-affinity chromatography ................................................................. 180
Figure 6.15: Cinnamate-4-hydroxylase activity of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr in a resting cell assay ................................ 181
Figure 6.16: SDS-PAGE and western blot analysis of Arabidopsis P450 73A5tr fused Arabidopsis ATR2tr or to Rhodococcus sp. RhF reductase in the soluble protein fraction .................................................. 182
Figure 6.17: SDS-PAGE and western blot analysis of Arabidopsis P450 73A5tr fused Arabidopsis ATR2tr or to Rhodococcus sp. RhF reductase in the insoluble protein fraction .................................................. 183
Figure 7.1: Schematic organisation of artificial A) plant-plant fusions (P450-ATR2tr) and B) plant bacterial (P450-RhF) fusions ................................. 193
Acknowledgements

At this point, I would like to say thank you to everyone who has helped me along the way to my PhD.

In particular, I would like to thank Prof. Neil C. Bruce and Dr. Gideon Grogan for great supervision, motivation and the opportunity for traveling around the world to attend the international P450 conferences, and for reminding me that plant P450s are a great challenge.

Thanks to Prof. Ian Graham and Dr. Gavin Thomas for the helpful advice at the training committee meetings.

Special thanks to all the members of the Bruce Group, especially Dr. Rosamond Jackson, Dr. Elizabeth Rylott, Dr. Astrid Lorenz and Dr. Joseph Bennett. Their kind help and endless support has helped me became a good lab member, an innovative report writer and creative presenter. Margaret, Helen, Chong, Andy, Will, Florian, Federico, Hazel, Mariya, Emily and Dana, I am grateful for the inspiring work environment and an unforgettable time here in York. Thank you very much!

Thanks to Prof. Mattheos Koffas for providing the IFS-CPR fusion, to Prof. Rodney Croteau for the CYP71D15 PM2-2 and Prof. Danièle Werck-Reichhart for the Saccharomyces cerevisiae strains WAT11 and WAT21, as well as for the plasmids pYeDP60-CYP89A9 and pYeDP60-CYP81D11.

Thanks also to the Centre of Excellence for Biocatalysis, Biotransformations and Biocatalytic Manufacture (CoEBio3) for their kind sponsorship.

Ganz lieb möchte ich mich auch bei meiner Familie und meinen Freunden bedanken. Durch eure Unterstützung habt ihr alle einen großen Anteil am Gelingen dieser Arbeit!

Meinem geliebten Ehemann, bestem Freund und größtem Kritiker Martin möchte ich für seine Liebe, Motivation und unermüdliche Unterstützung in all meinen Entscheidungen danken.
Author’s Declaration

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

1.1 Biotransformations in Industry

The amazing world of biotechnology was started by humans more than a thousand years ago by producing alcohol and vinegar by fermentation without even knowing about the presence of microorganisms. Later, living cells, such as yeast, bacteria, filamentous fungi and plants were deliberately used to improve the stability and taste of food. Industrial biotechnology, also called “white” biotechnology, is a relatively new definition and means the fabrication of products by using microorganisms or enzymes by reducing energy and waste for the creation of a more environmentally beneficial process.1 Over the last fifty years, many useful ‘biotransformation’ reactions have been identified and applied in industrial chemistry for processes including the hydrolysis and synthesis of ester and amide bonds, the hydrolysis of nitriles, the asymmetric reduction of prochiral ketones and the formation of chiral amines. Biotransformations have the advantage of converting substrates under relatively mild reaction conditions (neutral pH, room temperature, less activation energy) in comparison to conventional harsh chemical reactions. Moreover, enzymes are stereo- and regioselective and to recreate this selectivity is often a major challenge for synthetic and pharmaceutical chemistry using chemicatalysis. Therefore, enzymes are commercially important and often provide a more environmental friendly way for the synthesis of chemical compounds.

One particularly valuable class of enzymes are the oxygenase enzymes, which catalyse the introduction of oxygen into nonfunctionalised carbon skeletons. Oxygenations using conventional abiotic chemistry are energy consuming, often require either expensive catalysts and/or toxic reagents and are often neither regio- nor stereoselective. Hence there is now a substantial research effort into the discovery and application of enzymes that catalyse oxygenation reactions2, which has included peroxidases3, flavin dependent monooxygenases4, and heme-containing enzymes known as cytochromes P450. Amongst these enzymes exist biocatalysts capable of the hydroxylation of C-H bonds, heteroatoms such as
nitrogen and sulphur and many other useful enzymes that are distinguished by their ability to catalyse their reactions with high chemo, regio- and stereo-selectivity.

1.2 Cytochromes P450

Cytochromes P450 (P450s) belong to the enzyme class monooxygenases, which has the Enzyme Commission number (numeral classification of enzymes based on the reaction type) E.C.1.14.-.-. These heme-containing proteins are found in all kingdoms of life including humans, plants, insects, fungi, bacteria and also in viruses. In eukaryotes, they are mostly integral membrane bound, whereas prokaryotic P450 systems are more likely soluble and located in the cytoplasm. P450s require auxiliary reductases for the activation of molecular oxygen for their different reactions. These reductases transfer two electrons in single steps from the cofactor NAD(P)H to the heme of the P450.

1.2.1 History and Nomenclature of Cytochromes P450

Ronald W. Estabrook, one of the pioneers in P450 research, published a summary of the outstanding research in the discovery of the cytochromes P450 in 2003. It all began in 1949 with Betty and Jim Miller and their graduate student Gerald C. Mueller, who performed in vitro studies on the metabolism of methylated aminoazo dyes in rat liver homogenates, to understand the degradation of the carcinogen 4-dimethylaminoazobenzene. Later, Jim Gillette and Julius Axelrod paved the way for a better understanding of the drug (acetanilide) metabolism, which also gave important momentum to the discovery of the P450s. The latter worked on the demethylation of ephedrine and showed that NADPH and oxygen are necessary for the reaction in rabbit liver microsomes. Axelrod obtained the Nobel Prize in Physiology and Medicine (1970) together with Sir Bernard Katz and Ulf von Euler “for their discoveries concerning the humoral transmitters in the nerve terminals and the mechanism for their storage, release and inactivation”. Bernard B. Brodie, group leader of Jim Gillette and Julius Axelrod, published a classic paper on the hydroxylation of acetanilide and demethylation of monomethyl-4-aminoantipyrine in rabbit liver microsomes in
1955. The research in the area of the steroid hormone metabolism also contributed to the discovery of P450s, responsible for the C21 hydroxylation of progesterone in microsomes from the bovine adrenal cortex (Figure 1.14). Ryan and Engel showed in their studies the reversible inhibition of carbon monoxide by light for the first time. This observation led later to the discovery of P450 function as an oxidase using molecular oxygen and NADPH for their reactions. A breakthrough in P450 research – built on the knowledge and techniques found on cytochrome b5 by Britton Chance and G. Ron Williams – was the detection of a shift from 413 nm to 450 nm in the carbon monoxide-bound form of an unknown pigment from rat liver by Martin Klingenberg in 1958 (Figure 1.1).

Figure 1.1: Original figure of the carbon monoxide difference spectra present in rat liver microsomes (from Klingenberg 1958)

Enzymes of the superfamily cytochrome P450 received their name from Omura and Sato, who identified them as heme proteins in 1962. The P stands for “pigment” and the 450 for the characteristic absorption maximum at 450 nm (Soret peak) of the carbon monoxide bound form. This characteristic was also used to investigate a quantitative method for the detection of cytochromes
From this time on, there was a steady increase in the numbers of P450 publications in each year (Figure 1.2). Since 2010, the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) has shown that more than 2000 papers on P450s have been published each year.

Figure 1.2: Scientific articles published on National Center for Biotechnology Information (23th January 2012)

Today, there are more than 15000 P450 genes identified (Figure 1.3).

Figure 1.3: Schematic distribution of cytochromes P450 among the different kingdoms

One third of all known P450s belong to the plant kingdom and the P450 content in angiosperm genomes can reach up to 1%, whereas the typical vertebrate genome
contains less than 100 (e.g. humans with 57 and *Drosophila melanogaster* with 90 P450 genes). A few bacteria, such as the Gram-negative *Escherichia coli* were found to have no P450 gene sequence.

The first classification of 154 different genes of the P450 family from mammals, such as human, bovine, and rat as well as from yeast and bacteria (P450cam from *Pseudomonas putida*) was in 1991. Proteins sharing at least 40% of their amino acid sequence are classified into the same family and are assigned with the designation CYP for cytochrome P450 followed by a number (Figure 1.4) by Nebert and coworkers. The plant genes start at CYP71 through to CYP99 and continue again from CYP701-999. Sequences of the same enzyme family members with a similarity of more than 55% in their amino acid sequence were arranged into subfamilies such as CYP71A and CYP71B. There are some exceptions, especially in plants, where the classification is based on phylogenetic data and the organisation of the genes. The last number in the CYP name is unique to a single P450. Discoveries of new P450s resulted sometimes in modification of already named P450s. To monitor this, David Nelson (member of the Cytochrome P450 nomenclature committee) created a webpage with all identified P450s (http://drnelson.uthsc.edu/CytochromeP450.html).

### CYP51G1

![Diagram](http://drnelson.uthsc.edu/CytochromeP450.html)

Figure 1.4: Nomenclature of P450s

P450 and reductase form the P450 system and they are grouped in three main types.

The members of the eukaryotic P450 superfamily are associated with the membrane of the endoplasmatic reticulum and so called ‘microsomes type’. This system has a single flavin adenine dinucleotide / flavin mononucleotide (FAD/FMN) as the reductase with a flavoprotein P450 (Figure 1.5A). Both are synthesised by membrane-bound ribosomes and then anchored by an uncleavable
hydrophobic N-terminus (with around 20 amino acids of the N-terminus) into the membrane.\textsuperscript{30-32}

The second type (‘mitochondrial type’), uniquely found in animals but not in plants or fungi, consists of a P450 bound to the inner mitochondria membrane\textsuperscript{33} and a soluble FAD-containing flavoprotein with a ferredoxin-type iron-sulphur protein.\textsuperscript{34,35} However, some P450s have been reported to be present at the outer chloroplastic membrane, e.g. the Arabidopsis CYP74\textsuperscript{36,37}, CYP86B1\textsuperscript{38} and CYP701A3\textsuperscript{39} as well as an unnamed putative chloroplast P450 (AF107765)\textsuperscript{6} from \textit{Prunus dulcis} (almond).

The third type are the prokaryotic P450s (‘bacterial type’), where the P450 and the reductase parts are located solubly in the cytosol. Bacterial systems mostly utilise a ferredoxin reductase (FdR) combined with a ferredoxin (Fdx) as reductase part (Figure 1.5B) and generally use NADH as an electron donor.

![Figure 1.5: Schematic organisation of A the plant cytochrome and B the bacteria P450 systems](image)

The substrates for cytochromes P450 are mostly hydrophobic organic compounds.\textsuperscript{24} P450s have different functions and are involved in the biosynthesis of endogenous molecules as well as in the metabolism of many pharmaceuticals and xenobiotics. Across the kingdoms, P450s take part in a diverse range of processes: in prokaryotes, they participate in the assembly of antibiotics, in the catabolism of different carbon sources and in the metabolism of fatty acids.\textsuperscript{28} By contrast in eukaryotes, P450s are involved in the biosynthesis of membrane sterols. In animals, they are part of the biosynthesis of signal molecules and steroid hormones as well as vitamin D\textsubscript{3}.\textsuperscript{28} In fungi, they carry out important roles in the synthesis of mycotoxins and in the metabolism of lipids, which are used as carbon
source. Furthermore, in plants, they are found in the biosynthesis and catabolism of hormones and secondary plant compounds.

1.2.2 Structure and Function of Cytochromes P450

P450s between the different families share low sequence identity (often less than 20%). However, their structural organisation shows a strong conserved topology and three-dimensional fold. The number of crystal structures known for P450s is increasing rapidly since the first structure, of CYP101A1 (P450cam) from *Pseudomonas putida* was determined. Today, there are more than 70 crystal structures of P450s available on the Research Collaboratory for Structural Bioinformatics (RCSB) webpage (http://www.rcsb.org, 09/12/2011). The only solved crystal structure so far for Arabidopsis P450s is the allene oxide synthase (CYP74A1, Figure 1.6).

![Figure 1.6: Ribbon diagram of Arabidopsis allene oxide synthase (CYP74A1) crystal structure](image)

heme in grey, fatty acid substrate in green
The structure of allene oxide synthase in Figure 1.6 illustrates the general conserved topology of P450 structures. They contain approximately thirteen α-helices and four β-sheets ranged into two domains: the α domain is rich in helices and consists of the catalytic centre, a four-helix core (D, E, I and L) and another two helices (J and K), whereas the β domain has two β-sheets (β1 and β2) containing anti-parallel β-sheets, which are involved in the formation of the hydrophobic substrate channel (Figure 1.7).46

The comparison of two distantly related P450s, the eukaryotic CYP2C5 and the prokaryotic P450cam, revealed longer surface loops for the eukaryotic P450, whereas the catalytic centre and the L- and I-helices are the mostly conserved regions.47 The I-helix is involved in the substrate binding and the activation of the atomic oxygen.48

All P450s share the same catalytic centre, which consists of a heme B molecule (Figure 1.8) as a prosthetic group, which in almost all P450s is connected to the protein by a thiolate bond to a highly conserved cysteine residue. If this thiolate bond is missing, the characteristic absorbance maximum at 450 nm for the active conformational state shifted to 420 nm – the inactive species of P450s.
The hydrophobic heme iron environment (β-bulge segment) is a very conserved region, where the cysteine residue is stabilised by hydrogen bonds from peptide NH-groups assisting the regulation of the redox potential of the heme.\textsuperscript{49-51} A similar structure was also found in other enzyme groups such as nitric oxide synthase (NOS)\textsuperscript{52} and chloroperoxidase (CPO)\textsuperscript{53}.

Six substrate recognition sites (SRS) were identified by Gotoh using bioinformatics as dynamic protein regions opening and closing the substrate channel: SRS1 is on the B’helix, SRS2 and 3 are regions of the F- and G-helices, SRS4 is situated on the I-helix, SRS5 is the β4 hairpin and SRS6 the K-helix β1 connecting part.\textsuperscript{54,55}

Eukaryotic P450 have usually a longer hydrophobic N-terminus of 25-50 residues for the association with a membrane and the amino acid sequence is depended of the membrane target and not essential for the P450 function as shown for many human P450s.\textsuperscript{56-59}

Plant P450s often form metabolons with other proteins – a non-covalent complex of multible enzymes of a metabolic pathway. A substrate can be directly transferred into a product without producing high concentration of intermediates.\textsuperscript{60} One example for this phenomenon is the formation of a metabolon containing CYP79A1, CYP71E1, P450 reductase and the glycosyltransferase UGT85B1, which catalysis the reaction from tyrosine to dhurrin (a cyanogenic glucoside).\textsuperscript{61}
1.2.3 Reaction Mechanism

P450s are classified by their sequence similarity and not by their reaction type as most other enzymes. P450s show a wealth of reactions such as the cleavage of C-C-single bonds, C-N-double bonds, hydroxylations, sulfoxidations, epoxidations, deaminations, dehalogenations, peroxidations or N-oxide-reductions\textsuperscript{62,63} with more than thousands of accepted substrates.\textsuperscript{64} The ability to perform so many reactions may be due to the properties of the iron atom, which is able to change between different oxidation states.\textsuperscript{65}

The theory of the catalytic cycle of monooxygenases was developed in analogy with the reaction cycle of peroxidases, as P450s also use peroxide for monooxygenations.\textsuperscript{66,67} They require one oxygen and two electrons, which were usually donated in two single steps from the cofactor NAD(P)H to activate the oxygen, for the reduction of the heme iron\textsuperscript{28}.

\[
\text{RH} + \text{O}_2 + \text{NAD(P)H} + \text{H}^+ \rightarrow \text{ROH} + \text{H}_2\text{O} + \text{NAD(P)}^+ 
\]

At the start of the cycle the iron of the P450 has the oxidation state of +III and the sixth coordination centre is occupied by a molecule of water (Figure 1.9). The water ligand is displaced when a substrate is bound to the enzyme (1). As a result of this increase of the redox potential from a low spin state ($S = 1/2$) to high spin ($S = 5/2$), the enzyme can be reduced to P450-Fe(+II) by a single electron from the redox partner nicotinamide adenine dinucleotide (phosphate) (NAD(P)H) and the reductase (2).\textsuperscript{68,69} The result is a five times coordinate high-spin complex, which can bind different ligands such as molecular oxygen, carbon monoxide or cyanide (3). The oxygen-bound enzyme is similar to the superoxide-Fe(+III)-complex and is one of the detectable intermediates. The product of the following reduction through a single electron derived from the redox partners NAD(P)H and reductase (4) builds an unstable peroxo complex. After a protonation of the distal oxygen (5), it becomes a hydroperoxo intermediate (Compound 0), from which a water molecule is released after another protonation (6). The resulting product (Compound I) is unstable and was not isolated and characterised until recently (Rittle and Green 2010\textsuperscript{70}). Compound I is a high valent iron-oxo-complex with coordinated active oxygen responsible for the substrate conversion.\textsuperscript{70,71} Finally, the oxygen is transferred to the substrate (7) and a water molecule is initiated (8) so that the oxidised product and the initial state are formed.\textsuperscript{72-76}
The steps of the catalytic cycle are dependent on the substrate, electrons and protons. This avoids a consumption of reduction equivalents without oxygenation of the substrate, preventing futile cycling and stopping production of toxic superoxides. However, there are three major reactions (called “uncoupling”), which abort the catalytic cycle: The first is the autooxidation of the superoxide-Fe(+III)-complex by forming superoxide radicals. The second uncoupling, known as the peroxide shunt, separates hydrogen peroxide from the hydroperoxo intermediate. Some P450s use a short cut for their oxygenation reactions by introducing the oxygen from hydrogen peroxide into the substrate and using the peroxide shunt backwards, for example in fatty acid hydroxylation by CYP152A1 from *Bacillus subtilis* and CYP152B1 from *Sphingomonas paucimobilis*. The third reaction is called oxidase shunt, where the active oxygen species

---

**Figure 1.9: Reaction cycle of a P450 hydroxylation**

1. Substrate binding.
2. Reduction of the iron by NAD(P)H and reductase.
3. Oxygen inserting.
4. Reduction by NAD(P)H and reductase.
5. First protonation.
7. Substrate oxidation.
8. Substrate release.

The steps of the catalytic cycle are dependent on the substrate, electrons and protons. This avoids a consumption of reduction equivalents without oxygenation of the substrate, preventing futile cycling and stopping production of toxic superoxides. However, there are three major reactions (called “uncoupling”), which abort the catalytic cycle: The first is the autooxidation of the superoxide-Fe(+III)-complex by forming superoxide radicals. The second uncoupling, known as the peroxide shunt, separates hydrogen peroxide from the hydroperoxo intermediate. Some P450s use a short cut for their oxygenation reactions by introducing the oxygen from hydrogen peroxide into the substrate and using the peroxide shunt backwards, for example in fatty acid hydroxylation by CYP152A1 from *Bacillus subtilis* and CYP152B1 from *Sphingomonas paucimobilis*. The third reaction is called oxidase shunt, where the active oxygen species...
(Compound I) is protonated and reduced again by the dissociation of water instead of an insertion of the oxygen into the substrate.

1.2.4 Cytochromes P450 reactions

P450s provide a great variety of different regio- and stereospecific reactions, which make them interesting for industrial biocatalysis. Almost all P450s depend on a partner reductase and equimolar amounts of cofactors (NAD(P)H), which is a disadvantage when using these enzymes due to the increase in expense, however this problem can be avoided by performing whole cell experiments. The regeneration of NAD(P)H is also possible. For example, P450 BM3 (native fusion of a P450 and a reductase from *Bacillus megaterium*) can be coexpressed together with a glucose dehydrogenase. Glucose dehydrogenase converts glucose to gluconolactone in the process reducing NAD(P)⁺ to NAD(P)H, which is then used for the conversion of indole to indoxyl by P450 BM3. Indoxyl can dimerise to indigo, which is an important textile dye.  

Using the knowledge of computer modeling, genomics and proteomics, P450s have been applied in industry for the investigation of new drugs, medicine or xenobiotics.  

P450s with their wide range of substrate specificity catalyse key steps in many different pathways: In humans, they are mainly responsible for drug detoxification, in plants, for the activation or degradation of xenobiotics as well as for the biosynthesis of biochemical compounds and in bacteria, for the biodegradation of carbon sources.  

The variety of reactions catalysed by P450s extends from hydroxylation and oxidation, which are the most common P450 reactions, to alkylation, dealkylation, epoxidation, demethylation, aryl migration and many more (Figure 1.10). P450s catalysing hydroxylations, oxidations and epoxidations are an important alternative for industry, because chemical problems such as the use of chlorinated solvents and heavy metals as part of the catalyst are removed.
1.2.5 Interactions between Cytochromes P450 and reductases

Early studies of mammalian P450s and their reductases indicated the formation of a complex consisting of one P450 and one reductase molecule.\textsuperscript{89-93} Different research groups were studying the different P450 systems to understand the electron transfer between reductase and P450. In the microsomal P450 system, where the P450 and the reductase are both membrane associated, the electron transfer was reported to happen by random collision between P450 and reductase on the membrane\textsuperscript{94} or by a transient complex formation of the P450 and the reductase\textsuperscript{92,95,96}. Additionally, it was noticed that hydrophobic regions on the protein surface, such as the F-G loop, interact also with the membrane.\textsuperscript{97,98}

The number of mitochondrial P450 genes in animals is much smaller than for microsomal P450s. For example, seven mitochondrial P450s out of 57 are present in humans.\textsuperscript{35} Mitochondrial P450s lack the hydrophobic membrane anchor and therefore their association with the inner mitochondria membrane is weak compared to the membrane anchored microsomal P450s. The hydrophobic F-G loop (e. g. positions 219-237 for CYP27A1) is thought to be responsible for the
interaction between P450 and the inner mitochondrial membrane.\textsuperscript{99,100} The reductase part of the mitochondrial system contains two parts: the adrenodoxin and the adrenodoxin reductase and both interact via electrostatic interactions. Adrenodoxin, a soluble Fe\textsubscript{2}S\textsubscript{2} protein with high affinity to P450s, and the adrenodoxin reductase, are associated with the inner mitochondria membrane which enables the electron transfer between NADPH, reductase and P450.\textsuperscript{101-105} It was also demonstrated that mitochondrial and microsomal P450s and reductases can be interchangeable. A mitochondrial P450 was artificially anchored to the endoplasmic reticulum and shown to be supported by a microsomal reductase system as well as a microsomal P450 by the mitochondrial reductase.\textsuperscript{104,106,107} Bacterial P450 systems are soluble due to the fact that they are missing the hydrophobic membrane anchor. They are NAD(P)H dependent and consist usually of a P450, a ferredoxin and a ferredoxin reductase (Figure 1.5B). The genes for the bacterial P450 system are encoded on the chromosome or on plasmids, however, some bacteria such as \textit{E. coli} have no P450.

There is one example, where the reductase is not required for the P450 reaction: androstenedione (4-androstene-3-17-dione) is hydroxylated \textit{in vitro} by a P450 from rat liver in presence of NADPH and sodium periodate (NaIO\textsubscript{4})\textsuperscript{108}. Another exception is the soluble P450nor (nitric oxide reductase, CYP55A1) from \textit{Fusarium oxysporum}, which uses electrons directly from the cofactor NADH for the reduction of nitric oxide (NO) to nitrous oxide (N\textsubscript{2}O).\textsuperscript{109} Homologous of P450nor have been found in \textit{Cylindrocarpon tonkinense}\textsuperscript{110}, \textit{Trichosporom cutaneum}\textsuperscript{111,112} and \textit{Histoplasma capsulatum}\textsuperscript{113} possibly as a result of horizontal gene transfer\textsuperscript{114}.

1.2.6 Cytochrome P450 Fusion Systems

As the P450 reaction is typically dependent on the presence of single electrons derived from P450 reductases, co-expression is necessary to obtain effective P450 activity in recombinant systems.\textsuperscript{115-118} In Nature, the difficulty associated with using two separate enzymes is avoided by producing fusion P450-reductase enzymes, where usually the C-terminus of the P450 is fused to the N-terminus of the reductase. This phenomenon has only been found naturally to occur in bacteria and fungi so far.
The first self sufficient P450, where no additional enzymes are necessary, was the P450 BM3 (CYP102A1) from *B. megaterium* reported in 1986.\(^\text{119}\) The P450 (C-terminal) region of P450 BM3 is fused naturally into a single polypeptide chain to the cytochrome P450 reductase containing a FAD and a FMN. It has shown one of the highest turnover rates of all P450s, with about 17000 min\(^{-1}\) (when arachidonic acid is oxygenated\(^\text{120}\)), thus it is an extremely effective biocatalyst and this improved efficiency is one possible evolutionary reason for the fusion of P450 and reductase.\(^\text{119-122}\) A second advantage may be that the regulation of a P450 fusion system is easier than regulating expression of the proteins separately.\(^\text{123}\) However, the electron transfer between the three fusion domains of P450 BM3 has been shown to be mostly intermolecular by forming a dimer.\(^\text{124,125}\)

Several homologues of P450 BM3 from different bacteria have been identified by sequencing, such as CYP102A2 and CYP102A3 from *Bacillus subtilis*\(^\text{126}\), and the fusion proteins have been cloned and recombinantly expressed in *E. coli*.\(^\text{127-129}\)

The P450 BM3 belongs to the P450-diflavin reductase (CPR) fusion systems (Figure 1.11A) and showed significant homology to mammalian P450s and mammalian P450 reductases.\(^\text{130}\) This fusion type was imitated for the development of the first artificial fusion in 1987 containing the rat CYP1A1 and an NADPH cytochrome P450 reductase (from rat liver) expressed in *Saccharomyces cerevisiae*.\(^\text{116}\)

Further examples of natural fusion P450s type P450 BM3 are the fungal fatty acid hydroxylase P450foxy (CYP505) from *Fusarium oxysporum* and the Fum6p from *Fusarium verticillioides* (NCBI: *Gibberella moniliformis*).\(^\text{131-133}\)

Fungi have usually various P450s and one or multiple separate corresponding reductase. Additionally, putative fungal P450-CPR fusions have been found in all filamentous ascomycetes except *Coccidioides immitis* by analysing fungal genomes.\(^\text{134}\)
Figure 1.11: Schematic arrangement of natural P450-reductase fusion systems
A P450-BM3, a P450-diflavin reductase (CPR) fusion; B P450-RhF, a P450-phthalate dioxygenase reductase-like fusion; C MCCYP51FX, a P450-ferredoxin (Fdx) fusion with separate ferredoxin reductase (FdR) and D XplA/B, a P450-flavodoxin (Fldx) fusion with separate flavodoxin reductase XplB. Fusions are represented by overlapping spheres.

Through PCR-based screens another type of prokaryotic P450 fusion was found, where the P450 N-terminus was fused to a Phthalate Family Oxygenase Reductase (PFOR) containing an FMN and an iron-sulphur (2Fe-2S). One example belonging to this fusion type is the P450-RhF (CYP116B2) from Rhodococcus sp. (Figure 1.11B). The native substrate of this P450 fusion is not known, however it can catalyse the O-dealkylation of 7-ethoxycoumarin (Figure 1.12).
The fusion type MCCYP51FX (CYP51) from *Methylococcus capsulatus* with 14α-demethylase activity against lanasterol is fused with its C-terminus to a ferredoxin domain (Figure 1.11C).\(^{138,139}\)

Another class of P450 fusion identified first in the *Rhodococcus rhodochrous* strain 11Y contains the P450 XplA heme C-terminally fused to a flavodoxin domain and its separate flavodoxin reductase XplB (Figure 1.11D). XplA catalyses the first step in the degradation of the toxic explosive hexa-hydro-1,2,5-trinitro-1,3,5-triazine (RDX).\(^{140,141}\)

Bacteria are often used as host for the expression of P450 fusion proteins. For example, two mammalian cytochrome P450s (bovine adrenal P450 A17 and CYP4A1 from rat liver) separately fused to the flavoprotein domain of the rat NADPH cytochrome P450 reductase have been expressed in active forms in *E. coli*.\(^{142,143}\) This technique of producing artificial fusions was used in the same field of animal steroid hydrolases to study the human CYP93A4.\(^{144}\)

The FMN/Fe\(_2\)S\(_2\) reductase (RhF) from *Rhodococcus* sp. strain NCIMB 9784\(^{135,136}\), which is naturally fused to the CYP116B1, was expressed in *E. coli* Rosetta 2 (DE3) and was used to generate fusions with prokaryotic P450s such as P450 PikC (CYP107L1, involved in the last step of pikromycin biosynthesis) from *Streptomyces venezuelae*\(^{145}\), P450-cam (camphor-5-monooxygenase, CYP101) from *Pseudomonas putida*, XplA (CYP177A1) from *Rhodococcus rhodochrous* Y11, which degrades RDX; Noc4 and Noc10, two P450s from *Nocardia farcinica*, which can dealkylate 7-ethoxycoumarin and hydroxylate testosterone, respectively.\(^{146}\) The electron transfer between the different domains has been shown to be predominantly intramolecular for the artificial reductase-putidaredoxin-cytochrome P450cam triple fusion and the fusions P450cam-RhF-Red and XplAP450-RhF-Red.\(^{146,147}\)

---

**Figure 1.12: O-dealkylation of 7-ethoxycoumarin to 7-hydroxycoumarin by P450-RhF**

\[
\begin{align*}
\text{H}_3\text{C} & \quad \text{O} \quad \text{O} \quad \text{O} \\
\text{7-ethoxycoumarin} & \quad \text{P450-RhF} \\
\text{HO} & \quad \text{7-hydroxycoumarin}
\end{align*}
\]
1.3 Applications of Cytochromes P450 in biocatalysis

In the last few years, research in cytochromes P450 has expanded to include engineering of the biocatalysts for various applications. This has been reviewed in recent papers (Gillam 2008, Hlavica 2009, Grogan 2011, O’Reilly et al. 2011). For example, the self-sufficient P450 BM3 was modified at its substrate access channel with site-directed mutagenesis to increase the activity up to 40 fold towards the polycyclic aromatic hydrocarbons phenanthrene, fluoranthrene and pyrene.

1.3.1 Hydroxylation

The regio- and stereoselective oxidative activation of a C-H bond is a challenging problem for synthetic organic chemistry. Enzymes, especially P450s, are an environmental friendly solution for hydroxylations with high selectivity.

P450-catalysed hydroxylation occurs by an activation of a C-H bond followed by the insertion of an oxygen atom and finally the formation of the corresponding alcohol. The first solubly expressed P450 derived from *Rhizobium* with azoreductase activity was published in 1967. One year later, another solubly expressed P450, the methylene hydroxylase (P450$_{cam}$) converting camphor to 5-exo-hydroxycamphor from *Pseudomonas putida* was reported (Figure 1.13). Another organism, *Rhodococcus* sp. NCIMB 9784, is also able to use camphor as only carbon source. The P450 responsible for this metabolic activity is called P450camr, hydroxylating camphor to 6-endo-hydroxycamphor.

![Figure 1.13: Hydroxylation of camphor by P450cam (Pseudomonas putida) to 5-exo-hydroxycamphor or P450camr (Rhodococcus sp.) to 6-endo-hydroxycamphor](image)

Generally, bacteria use P450 hydroxylations as the first steps of the biodegradation of carbon sources, for example the degradation of cineole by
*Citrobacter braakii*[^159], or for the biosynthesis of secondary metabolites, such as antibiotic or neurotoxin by *Streptomyces* sp.[^160-162]. Another prokaryotic P450, the CYP153A6 from *Mycobacterium* sp. HXN-1500 hydroxylates alkanes with a chain-length between C6 and C11 to primary alkanols with a large regiospecificity of over 95%.[^163]

The alkane hydroxylase of the CYP153 family from *Mycobacterium* sp. catalyses the conversion of limonene to perillyl alcohol, an anticancer drug. The hydroxylase can be recombinantly expressed in *Pseudomonas putida* in large bioreactors to produce sufficient enzyme for pharmaceutical applications.[^164]

One of the first discovered P450 with hydroxylation activity was the human CYP21A2, which converts progesterone to deoxycorticosterone (Figure 1.14).[^14]

![Figure 1.14: Hydroxylation of progesterone to deoxycorticosterone by CYP21A2](image)

Other human P450s have been shown to be involved in steroid metabolism such as CYP2A19, which hydroxylates progesterone to 21-hydroxyprogesterone as the main product forming 16α- and 17α-hydroxyprogesterone as minor products. Furthermore, it was shown that CYP2A19 converted testosterone to androstenedione with low amounts of 2ß-, 6ß- and 16ß-hydroxy testosterone.[^165]

In the plant *Arabidopsis thaliana*, different cytochromes catalyse hydroxylations involved in different biosyntheses. One of the first P450s examined was the *Arabidopsis* cinnamic acid 4-hydroxylase (CYP73A5), which is involved with the 3’-hydroxylase (CYP98A3) and the 5-hydroxylase (CYP84A1) in the phenylpropanoid pathway and the biosynthesis of lignin.[^166-168]

The metabolism of brassinolide involves a 22α-hydroxylase CYP90B1 called DWF4 catalysing the conversion of 6-oxo-campestanol to cathasterone.[^169,170]
(Figure 1.15) and the 23α-hydroxylase CYP90A1 called cpd for 6-oxo-cathasterone\textsuperscript{171,172}. On the other hand, P450s, for example the 26-hydroxylase CYP72B1, also take part in the degradation of brassinolide\textsuperscript{173}.

Figure 1.15: Hydroxylation of 6-oxo-campestanol to cathasterone by Arabidopsis CYP90B1

1.3.2 Oxidation

P450s are monooxygenases and oxidise the substrate following the equation:

\[
\text{RH} + \text{O}_2 + 2\text{H}^+ + 2e^- \rightarrow \text{ROH} + \text{H}_2\text{O}
\]

Oxidases, dioxygenases and peroxidases catalyse similar reactions, however P450s are special, because they are able to introduce oxygen into double bonds, allylic positions and also non-activated C-H bonds\textsuperscript{174}.

Another typical P450 catalysed reaction is the oxidation of heteroatoms, for example nitrogen in tertiary amines to N-oxides or sulphur in thioethers to sulfoxides and sulfones (Figure 1.16).

Figure 1.16: P450 dependent heteroatom-oxidations of tertiary amines to N-oxides and thioether to sulfoxides or sulfones
An enzymatic oxidation of methyl groups on aromatic heterocycles, such as pyridines, to the corresponding monocarboxylic acids was investigated by fermentation of *Pseudomonas putida* on industrial scale by Lonza AG (Figure 1.17).\textsuperscript{175} Pyridine-3-carboxylic acid (niacin) is the water-soluble vitamin B3 used in animal feed supplementation and medicine.\textsuperscript{176,177}

![Figure 1.17: P450 catalysed oxidation of 3-methylpyridine to pyridine-3-carboxylic acid](image)

P450s catalyse also the epoxidation of double bonds, for example the *Arabidopsis thaliana* (*Arabidopsis*) CYP77A4 can epoxidise unsaturated C18 fatty acids.\textsuperscript{178} Epoxidations result in very reactive versatile epoxide species by transferring the P450 ferryl-oxygen to the alkene, which builds a radical intermediate followed by the epoxide formation (Figure 1.18).

![Figure 1.18: P450 catalysed epoxidation of an alkene](image)

Epoxide can easily react with complex macromolecules useful for the synthesis of polymers\textsuperscript{179} and pharmaceuticals\textsuperscript{83} for industrial applications.
1.4 Plant Cytochromes P450

1.4.1 Cytochromes P450s in *Arabidopsis thaliana*

*Arabidopsis* is a much used model plant in biology, because of its fast growth cycle (48 days until the first siliquae)\(^1\) and ease of transformation. The whole genome is sequenced, and it is known that this plant has over 25000 genes, which code proteins from 11000 families.\(^2\) *Arabidopsis* has 244 P450 genes and 28 pseudogenes\(^3\) and represents a fantastic opportunity for a high number of different biocatalysts to be discovered.

The phylogenetic tree (adapted from Bak *et al.*2011\(^4\), Figure 1.19) shows the evolutionary relationships between the different P450s in *Arabidopsis*. 

---

\(^1\) Bak *et al.* (2011)

\(^2\) Arabidopsis has over 25000 genes, which code proteins from 11000 families.

\(^3\) Arabidopsis has 244 P450 genes and 28 pseudogenes.

\(^4\) Adapted from Bak *et al.* (2011)
Figure 1.19: Phyllogenetic tree of all cytochromes P450 from Arabidopsis

The 67 Arabidopsis P450s with known function are listed in Table 1.1. They are mainly involved in secondary metabolism, such as in the biosynthesis of hormones, pigments, flavours and for the production of complex macromolecules such as cutin, lignin and suberin or for defense compounds. Approximately 20 P450 contain a rich N-terminal Serine-Threonine sequence indicating a chloroplast signal.
Table 1.1: Arabidopsis P450s with known function

<table>
<thead>
<tr>
<th>P450 (gene locus)</th>
<th>reaction</th>
<th>pathway</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>51G1 (At1g11680)</td>
<td>obtusifoliol 14α-demethylase</td>
<td>sterols/steroids</td>
<td>183,184</td>
</tr>
<tr>
<td>71A12 (At2g30750)</td>
<td>hydroxylation and N-demethylation of pyrazoxyfen</td>
<td>metabolism of herbicide pyrazoxyfen, camalexin biosynthesis in roots</td>
<td>185,186</td>
</tr>
<tr>
<td>71A13 (At2g307700)</td>
<td>dehydration of indole acetaldoxime to indole-3-acetonitrile</td>
<td>camalexin biosynthesis</td>
<td>187</td>
</tr>
<tr>
<td>71B7 (At1g13110)</td>
<td>deethylation of 7-ethoxycoumarin (with cumene hydroperoxide as electron donor)</td>
<td></td>
<td>188</td>
</tr>
<tr>
<td>71B15 (At3g26830)</td>
<td>conversion of s-dihydrocamalexic acid to camalexin</td>
<td>camalexin</td>
<td>189,190</td>
</tr>
<tr>
<td>73A5 (At2g30490)</td>
<td>4-hydroxylation of t-cinnamic acid to p-coumaric acid; hydroxylation of cinnamic acid analogs</td>
<td>phenylpropanoid pathway, lignin biosynthesis</td>
<td>166,191,192</td>
</tr>
<tr>
<td>74A1 (At5g42650)</td>
<td>allene oxide synthase for linoleic acid hydroperoxide and linolenic acid hydroperoxide</td>
<td>jasmonate acid biosynthesis</td>
<td>37,45,193-196</td>
</tr>
<tr>
<td>74B2 (At4g15440)</td>
<td>hydroperoxide lyase for linoleic acid hydroperoxide and linolenic acid hydroperoxide</td>
<td>oxylipin pathway, jasmonate biosynthesis</td>
<td>197-199</td>
</tr>
<tr>
<td>75B1 (At5g07990)</td>
<td>3'-hydroxylation of narigenin and dihydrokaempferol</td>
<td>phenylpropanoid pathway, flavonoid biosynthesis</td>
<td>200</td>
</tr>
<tr>
<td>76C1 (At2g45560)</td>
<td>10-hydroxylation of geraniol</td>
<td>terpenoid indole alkaloid biosynthesis</td>
<td>201</td>
</tr>
<tr>
<td>77A4 (At5g04660)</td>
<td>epoxidation of unsaturated C_{18} fatty acids (oleic, linoleic, α-linolenic acid)</td>
<td>fatty acid metabolism</td>
<td>178</td>
</tr>
<tr>
<td>77A6 (At3g10570)</td>
<td>in-chain 10-hydroxylation of 16-hydroxypalmitate</td>
<td>cutin biosynthesis</td>
<td>202</td>
</tr>
<tr>
<td>79A2 (At5g05260)</td>
<td>conversion of phenylalanine to oxime</td>
<td>benzyllglucosinolate biosynthesis</td>
<td>203</td>
</tr>
<tr>
<td>79B2 (At4g39950)</td>
<td>conversion of tryptophan and tryptophan analogs to oxime</td>
<td>indole glucosinolate biosynthesis, camalexin biosynthesis, auxin biosynthesis</td>
<td>204-209</td>
</tr>
<tr>
<td>79B3 (At2g22330)</td>
<td>conversion of tryptophan and tryptophan analogs to oxime</td>
<td>indole glucosinolate biosynthesis, camalexin biosynthesis, auxin biosynthesis</td>
<td>204-209</td>
</tr>
<tr>
<td>79F1 (At1g16410)</td>
<td>N-hydroxylation of homo-tetrahomomethionine (n = 3 to 6) (short-chain methionine derivatives to their aldoximes)</td>
<td>biosynthesis of aliphatic glucosinolates</td>
<td>210-213</td>
</tr>
</tbody>
</table>
### Chapter 1 - Introduction

<table>
<thead>
<tr>
<th>P450 (gene locus)</th>
<th>reaction</th>
<th>pathway</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>79F2 (At1g16400)</td>
<td>N-hydroxylation of long chain penta and hexahomomethionine to their aldoxime</td>
<td>biosynthesis of aliphatic glucosinolates</td>
<td>211-213</td>
</tr>
<tr>
<td>81F2 (At5g57220)</td>
<td>4-hydroxylation of indole-3-ylmethyl to 4-hydroxy-indole-3-ylmethyl glucosinolate</td>
<td>indole glucosinolate biosynthesis</td>
<td>214-216</td>
</tr>
<tr>
<td>81F4 (At4g37410)</td>
<td>putative P450</td>
<td>possibly involved in the indole glucosinolate biosynthesis</td>
<td>217</td>
</tr>
<tr>
<td>82C2 (At4g31970)</td>
<td>5-hydroxylation of 8-methoxypsoralen to 5-hydroxy-8-methoxypsoralen</td>
<td>metabolism of tryptophan-derived secondary metabolites, possibly involved in jasmonic acid induced indole glucosinolates</td>
<td>218,219</td>
</tr>
<tr>
<td>82C4 (At4g31940)</td>
<td>5-hydroxylation of 8-methoxypsoralen to 5-hydroxy-8-methoxypsoralen</td>
<td>Fe deficiency response, possibly through an IDE1-like mediated pathway</td>
<td>218,220</td>
</tr>
<tr>
<td>82G1 (At3g25180)</td>
<td>(E,E)-geranyllinalool and the sesquiterpenoid (E)-nerolidol into the acyclic volatile C_{16}-homoterpenes 4,8, 12-trimethyltdec-1,3,7, 11-tetraene (TMTT) and the C_{11}-homoterpenes 4,8-dimethyl-1,3,7-nonathene (DMNT), respectively</td>
<td></td>
<td>221</td>
</tr>
<tr>
<td>83A1 (At4g13770)</td>
<td>oxidation of methionine-derived oximes oxidation of p-hydroxyphenyl-acetaldoxime, indole-3-acetaldoxime, conversion of aldoximes to thiohydroximates</td>
<td>biosynthesis of aliphatic glucosinolates</td>
<td>222-224</td>
</tr>
<tr>
<td>83B1 (At4g31500)</td>
<td>oxidation of indole-3-acetaldoxime</td>
<td>biosynthesis of indole glucosinolates</td>
<td>222,223,225</td>
</tr>
<tr>
<td>84A1 (At4g36220)</td>
<td>5-hydroxylation of coniferaldehyde, coniferyl alcohol and ferulic acid</td>
<td>phenylpropanoid pathway, biosynthesis of lignin</td>
<td>167,226,227</td>
</tr>
<tr>
<td>84A4 (At5g04330)</td>
<td>putative ferulate-5-hydroxylase</td>
<td></td>
<td>228,229</td>
</tr>
<tr>
<td>85A1 (At5g38970)</td>
<td>C6-oxidase for 6-deoxycastasterone to castasterone, other steroids conversion of castasterone to brassinolide</td>
<td>biosynthesis of brassinolide</td>
<td>230-234</td>
</tr>
<tr>
<td>85A2 (At3g30180)</td>
<td>C6-oxidase for 6-deoxycastasterone to castasterone, other steroids conversion of castasterone to brassinolide</td>
<td>biosynthesis of brassinolide</td>
<td>230-233,235,236</td>
</tr>
<tr>
<td>86A1 (At5g58860)</td>
<td>ω-hydroxylation of saturated and unsaturated fatty acid metabolism, biosynthesis of cutin,</td>
<td></td>
<td>237-240</td>
</tr>
<tr>
<td>P450 (gene locus)</td>
<td>reaction</td>
<td>pathway</td>
<td>references</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>86A2 (At4g00360)</td>
<td>ω-hydroxylation of saturated and unsaturated C12 to C18 fatty acids</td>
<td>fatty acid metabolism, biosynthesis of cutin</td>
<td>238,241,242</td>
</tr>
<tr>
<td>86A4 (At1g01600)</td>
<td>ω-hydroxylation of saturated and unsaturated C12 to C18 fatty acids</td>
<td>fatty acid metabolism, biosynthesis of cutin</td>
<td>202,238,242,243</td>
</tr>
<tr>
<td>86A7 (At1g63710)</td>
<td>ω-hydroxylation of lauric acid</td>
<td>fatty acid metabolism, biosynthesis of cutin</td>
<td>238,242</td>
</tr>
<tr>
<td>86A8 (At2g45970)</td>
<td>ω-hydroxylation of saturated and unsaturated C12 to C18 fatty acids</td>
<td>fatty acid metabolism, biosynthesis of cutin</td>
<td>238,242,244</td>
</tr>
<tr>
<td>86B1 (At5g23190)</td>
<td>ω-hydroxylation for long chain fatty acid (C22 and C24)</td>
<td>biosynthesis of suberin, fatty acid metabolism, polyester monomer biosynthesis</td>
<td>38,245</td>
</tr>
<tr>
<td>86C3 (At1g13140)</td>
<td>hydroxylation of fatty acids (C12, C14, C14:1, C16)</td>
<td>fatty acid metabolism</td>
<td>246</td>
</tr>
<tr>
<td>88A3 (At1g05160)</td>
<td>oxidation of ent-kaurenoic acid to gibberilin A12 in three steps</td>
<td>biosynthesis of gibberellins</td>
<td>39,247</td>
</tr>
<tr>
<td>88A4 (At2g32440)</td>
<td>oxidation of ent-kaurenoic acid to gibberilin A12 in three steps</td>
<td>biosynthesis of gibberellins</td>
<td>39,247</td>
</tr>
<tr>
<td>90A1 (At5g05690)</td>
<td>23α-hydroxylation of steroids</td>
<td>biosynthesis of brassinolide</td>
<td>171,172,248,249</td>
</tr>
<tr>
<td>90B1 (At3g50660)</td>
<td>22α-hydroxylation of C27, C28, C29 sterols</td>
<td>biosynthesis of brassinolide</td>
<td>169,249-251</td>
</tr>
<tr>
<td>90C1 (At4g36380)</td>
<td>C23-hydroxylation of sterols (typhasterol to castasterone)</td>
<td>biosynthesis of brassinolide</td>
<td>252-255</td>
</tr>
<tr>
<td>90D1 (At3g13730)</td>
<td>C23-hydroxylation of sterols</td>
<td>biosynthesis of brassinolide</td>
<td>253-255</td>
</tr>
<tr>
<td>94B3 (At3g48520)</td>
<td>12-hydroxylation of jasmonoyl-L-isoleucine</td>
<td>oxidative catabolism of jasmonate, jasmonate mediated signaling pathway</td>
<td>256,257</td>
</tr>
<tr>
<td>94C1 (At2g27690)</td>
<td>ω-hydroxylation of saturated and unsaturated C12 to C18 fatty acids and hydroxylation of ω-hydroxy fatty acid into dicarboxylic fatty acid</td>
<td>fatty acid metabolism</td>
<td>240,258</td>
</tr>
<tr>
<td>96A15 (At1g57750)</td>
<td>hydroxylation of midchain alkane to alcohols and second hydroxylation to ketone</td>
<td>biosynthesis of wax</td>
<td>259</td>
</tr>
<tr>
<td>97A3 (At1g31800)</td>
<td>β-hydroxylation of carotenone</td>
<td>carotenoid metabolism</td>
<td>260,261</td>
</tr>
<tr>
<td>97B3 (At4g15110)</td>
<td>β-hydroxylation of β-carotene to zeaxanthin via β-cryptoxanthin</td>
<td>carotenoid metabolism</td>
<td>262</td>
</tr>
<tr>
<td>97C1 (At3g53130)</td>
<td>ε-hydroxylation of β,ε-carotene</td>
<td>carotenoid metabolism</td>
<td>260,263</td>
</tr>
<tr>
<td>98A3 (At2g40890)</td>
<td>3-hydroxylation of p-coumarate to caffeic acid</td>
<td>phenylpropanoid pathway, biosynthesis of lignin</td>
<td>168,264-267</td>
</tr>
</tbody>
</table>
## Chapter 1 - Introduction

<table>
<thead>
<tr>
<th>P450 (gene locus)</th>
<th>reaction</th>
<th>pathway</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>and 3-hydroxylation of coumaroyl-esters (shikimate and quinate esters)</td>
<td>monomers and soluble phenolics</td>
<td></td>
</tr>
<tr>
<td>98A8 (At1g74540)</td>
<td><em>meta</em>-hydroxylation of the three triferuloylspermidine phenolic rings, oxygenation of resveratrol</td>
<td>alternative phenolic pathway, pollen development</td>
<td>268,269</td>
</tr>
<tr>
<td>98A9 (At1g74550)</td>
<td><em>meta</em>-hydroxylation of the three triferuloylspermidine phenolic rings, oxygenation of resveratrol</td>
<td>alternative phenolic pathway, pollen development</td>
<td>268,269</td>
</tr>
<tr>
<td>701A3 (At5g25900)</td>
<td>oxidation of <em>ent</em>-kaurene to <em>ent</em>-kaurenoic acid in three steps</td>
<td>biosynthesis of gibberellin</td>
<td></td>
</tr>
<tr>
<td>703A2 (At1g01280)</td>
<td>in-chain monohydroxylation of saturated fatty acids (C10-C16)</td>
<td>biosynthesis of sporopollenin, pollen development</td>
<td>274</td>
</tr>
<tr>
<td>704B1 (At1g69500)</td>
<td><em>ω</em>-hydroxylation of saturated, unsaturated and epoxy C16 and C18 fatty acids</td>
<td>biosynthesis of sporopollenin, pollen development</td>
<td></td>
</tr>
<tr>
<td>705A5 (At5g47990)</td>
<td>conversion of thalianol to desaturated thalianol</td>
<td>thalianol metabolic pathway</td>
<td>277</td>
</tr>
<tr>
<td>707A1 (At4g19230)</td>
<td>8-hydroxylation of abscisic acid</td>
<td>abscisic acid catabolism</td>
<td>278,280</td>
</tr>
<tr>
<td>707A2 (At2g29090)</td>
<td>8-hydroxylation of abscisic acid</td>
<td>abscisic acid catabolism</td>
<td>278,281</td>
</tr>
<tr>
<td>707A3 (At5g45340)</td>
<td>8-hydroxylation of abscisic acid</td>
<td>abscisic acid catabolism</td>
<td>278,279,281-286</td>
</tr>
<tr>
<td>707A4 (At3g19270)</td>
<td>8-hydroxylation of abscisic acid</td>
<td>abscisic acid catabolism</td>
<td>278,279</td>
</tr>
<tr>
<td>708A2 (At5g48000)</td>
<td>hydroxylation of thalianol to thalianol</td>
<td>thalianol metabolic pathway</td>
<td>277</td>
</tr>
<tr>
<td>710A1 (At2g34500)</td>
<td>C22-desaturation of β-sitosterol to stigmasterol</td>
<td>biosynthesis of sterols</td>
<td>287,289</td>
</tr>
<tr>
<td>710A2 (At2g34490)</td>
<td>C22-desaturation of β-sitosterol and 24-epi-campesterol to stigmasterol and brassicasterol</td>
<td>biosynthesis of sterols</td>
<td>287</td>
</tr>
<tr>
<td>710A4 (At2g28860)</td>
<td>C22-desaturation of β-sitosterol to stigmasterol</td>
<td>biosynthesis of sterols</td>
<td>288</td>
</tr>
<tr>
<td>714A1 (At5g24910)</td>
<td>C13-hydroxylation of <em>ent</em>-kaurenoic acid to steviol</td>
<td>gibberillin catabolism</td>
<td>277,290</td>
</tr>
<tr>
<td>724A1 (At5g14400)</td>
<td>22α-hydroxylation of brassinostearoid</td>
<td>biosynthesis of brassinolide</td>
<td>291</td>
</tr>
<tr>
<td>734A1 (At2g26710)</td>
<td>initially classified as CYP72B</td>
<td>brassinolide catabolism</td>
<td>173,292,293</td>
</tr>
<tr>
<td>735A1 (At5g38450)</td>
<td><em>trans</em>-hydroxylation of isopentenyladenine, tri/di/monophosphates</td>
<td>biosynthesis of zeatin, cytokinin metabolism</td>
<td>294</td>
</tr>
<tr>
<td>735A2 (At1g67110)</td>
<td><em>trans</em>-hydroxylation of isopentenyladenine, tri/di/monophosphates</td>
<td>biosynthesis of zeatin, cytokinin metabolism</td>
<td>294</td>
</tr>
</tbody>
</table>
The CYP51 is the most conserved P450 family across all kingdoms and counts as an eukaryotic P450 of prokaryotic origin. CYP51 is involved in the biosynthesis of sterols, e.g. cholesterol in animals, obtusifoliol in plants and ergosterol in fungi by catalysing 14α-demethylation (Figure 1.20). In plants, the methyl group of carbon number 14 is hydroxylated to 14α-carboxyalcohol and then converted to 14α-carboxyaldehyde. In the last step a double bond is introduced by an oxidative deformylation with separation of formic acid.

Figure 1.20: 14α-demethylation of obtusifoliol by plant CYP51

75% of the Arabidopsis P450s are still with unknown function, which shows that there is a greater research potential for finding new metabolic pathways and reactions in plants.

1.4.2 Fusions of plant Cytochromes P450

In Nature there exists no native plant P450 fusion so far. There are now a few examples where plant P450s (C-terminus) fused to plant P450 reductase have been expressed as active enzymes in E. coli (Table 1.2). The first artificial plant fusion contained the cinnamate-4-hydroxylase (CYP73A4) and the P450 reductase (both enzymes from C. roseus, Madagascar periwinkle) and was engineered by Prof. Schröder's research group in 1995. They found activity in the insoluble protein fraction after expression in E. coli. A few years later, they determined that a fusion of a P450 and the reductase from two different plants (Petunia hybrida CYP75 fused to C. roseus reductases) resulted in 50% lower specific activity than when the two enzymes are both from C. roseus. Moreover, they demonstrated that their strategy is an alternative technique for expression of
functional plant P450s and engineered another two fusions of CYP71D12 (tabersonine-16-hydroxylase from C. roseus) and CYP72A1 (secologanin synthase from C. roseus) fused to the reductase of C. roseus. Tabersonine-16-hydroxylase catalyses the first reaction in the pathway of vinblastine and vincristine, both bisindoles that are important for the pharmaceutical industry where they are used as treatments for leukaemia. Secologanin synthase converts loganin to secologanin, which is a precursor of terpenoid indole alkaloids in various plants important for development of drugs in cancer treatment. Another artificial plant fusion made of CYP71B1 from Thlaspi arvensae (field penny cress) and the C. roseus reductases was found to be active for the conversion of benzo(a)pyrene to 3-hydroxybenzo(a)pyrene in the membrane fraction after expression in E. coli.
<table>
<thead>
<tr>
<th>P450 from</th>
<th>Reductase from</th>
<th>Reference</th>
<th>Linker sequence (amino acid)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thlaspi arvensae (field pennycress)</td>
<td>Catharanthus roseus</td>
<td>Lamb et al. 1998&lt;sup&gt;305&lt;/sup&gt;</td>
<td>ST</td>
</tr>
<tr>
<td>Catharanthus roseus</td>
<td>Catharanthus roseus</td>
<td>Schröder et al. 1999&lt;sup&gt;301&lt;/sup&gt;</td>
<td>ST</td>
</tr>
<tr>
<td>Catharanthus roseus</td>
<td>Catharanthus roseus</td>
<td>Imler et al. 2000&lt;sup&gt;302&lt;/sup&gt;</td>
<td>ST</td>
</tr>
<tr>
<td>Catharanthus roseus</td>
<td>Catharanthus roseus</td>
<td>Houze et al. 1995&lt;sup&gt;309&lt;/sup&gt;</td>
<td>ST</td>
</tr>
<tr>
<td>Catharanthus roseus</td>
<td>Catharanthus roseus</td>
<td>Kaltenbach et al. 1999&lt;sup&gt;300&lt;/sup&gt;</td>
<td>ST</td>
</tr>
<tr>
<td>Catharanthus roseus</td>
<td>Catharanthus roseus</td>
<td>Kaltenbach et al. 1999&lt;sup&gt;300&lt;/sup&gt;</td>
<td>ST</td>
</tr>
<tr>
<td>Petunia hybrid (petunia)</td>
<td>Glycine max (soy bean)</td>
<td>Leonard &amp; Koffas 2007&lt;sup&gt;306&lt;/sup&gt;</td>
<td>GST</td>
</tr>
<tr>
<td>Oryza sativa (rice)</td>
<td>Trifolium pratense (red clover)</td>
<td>Kim et al. 2009&lt;sup&gt;307&lt;/sup&gt;</td>
<td>GST</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P450</th>
<th>Table 1.2: Plant fusion P450s heterologously expressed in E. coli</th>
</tr>
</thead>
<tbody>
<tr>
<td>P450</td>
<td>CYP71B1</td>
</tr>
<tr>
<td></td>
<td>CYP71D2 (tabersonine 16-hydroxylase)</td>
</tr>
<tr>
<td></td>
<td>CYP72A1 (secoisolariciresinol)</td>
</tr>
<tr>
<td></td>
<td>CYP73A4 (cinnamate 4'-hydroxylase)</td>
</tr>
<tr>
<td></td>
<td>CYP75 (flavonoid 3',5'-hydroxylase)</td>
</tr>
<tr>
<td></td>
<td>IPS (isoflavone synthase)</td>
</tr>
<tr>
<td></td>
<td>isoflavone synthase</td>
</tr>
</tbody>
</table>
More recent publications have specially focused on fusions containing the
isoflavone synthase I (IFS) from *Glycine max* (soy bean) and isoflavone synthase
from red clover.\textsuperscript{306,307} The artificial plant fusion system IFS-CPR (reductase CPR
from *C. roseus*) was created by Koffas’s research group by imitating the powerful
natural fusion P450 BM3 from *B. megaterium*.\textsuperscript{306} The IFS converts naringenin
and liquiritigenin to genistein and daidzein, respectively (Figure 1.21).

![Conversion of the flavonones liquiritigenin and naringenin to the
isoflavones daidzein and genistein, respectively, by the P450 isoflavone synthase I](image)

The reaction mechanism catalysed by IFS starts with an oxidation in the presence
of NADPH+H\textsuperscript{+} and molecular oxygen to a radical (1, Figure 1.22), followed by an
aryl ring migration (2) and forming a 2,7,4′-trihydroxyisoflavanone (3). The last
step is a dehydration (4), which yields genistein and is not performed by the P450.
It is possible that a specific dehydratase is involved or that the water molecule is
eliminated spontaneously.\textsuperscript{308-310}
Another artificial plant IFS-fusion for expression in \textit{E. coli} BL21 (DE3) was developed by Kim \textit{et al.} 2009\cite{307}. Therefore, the IFS from red clover was fused to the rice reductase (rcIFS-riceCPR), deleting the membrane anchors of \textit{rcIFS} (63 nucleotides) and \textit{riceCPR} (49 nucleotides), and activity has been demonstrated in a growing cell assay.\cite{307}

Isoflavones belong to the group of isoflavonoids and are predominantly produced as secondary metabolites in legumes.\cite{311} Isoflavones have a broad bioactivity spectrum: inhibition of tyrosine kinases and enzymes in steroid biosynthesis, antioxidant activity, prevention of carcinogenesis and inhibition of metastasis. In Asia, where soy bean (rich of isoflavones) is eaten much more than in western countries, a reduced frequency of heart diseases, osteoporosis, menopausal symptoms, breast and prostate cancer was recorded.\cite{312-317}

### 1.5 Recombinant Expression Systems

There are a number of different expression systems available. The two least expensive and most commonly used systems are the bacterium \textit{E. coli} (which has no endogenous P450 or P450 reductase genes) and the yeast \textit{Saccharomyces cerevisiae} (containing three P450 and one associated NAD(P)H P450 reductase gene).\cite{28,318} Furthermore, insect cells, another heterologous eukaryotic expression
system, are often used for the expression of plant P450s together with the insect NADPH P450 reductase.\textsuperscript{319} Duan and Schuler (2006)\textsuperscript{319} give a useful summary of the expression of Arabidopsis P450s in different hosts.

1.5.1 Expression of P450s in \textit{Escherichia coli}

The great advantage of bacterial expression over yeast or insect cells is the relatively fast growth rates, inexpensive cost and ease of cultivation. However, expression of unmodified plant P450s in \textit{E. coli} can be complicated due to differences in codon preference and because eukaryotic P450s are mostly membrane bound and bacteria do not have such a developed membrane system. Expression of plant P450s in \textit{E. coli} is therefore challenging and it is often difficult to obtain soluble protein.\textsuperscript{320,321} If the P450 is only anchored in the membrane, with the majority of the enzyme, including the catalytic centre in the cytosol, then removal of the hydrophobic membrane bound part may result in soluble enzyme. Studies on mammalian P450s have shown activity towards their specific substrates after the deletion of around 2-20 of the hydrophobic residues from the N-terminus.\textsuperscript{57,59,322} Additionally, the microsomal human P450 2C3 could be expressed solubly after the modification of the N-terminal hydrophobic membrane bound segment.\textsuperscript{58}

Many plant P450s have been expressed in \textit{E. coli}, none of them soluble, for biochemistry studies. The two Arabidopsis CYP79F1 and CYP79F2, both involved in the biosynthesis of aliphatic glucosinolates have been characterised after recombinant expression in \textit{E. coli}.\textsuperscript{210,212} Artificial P450 fusion proteins have been expressed also in \textit{E. coli} (more information in Chapter 1.4.2, page 43).

1.5.2 Expression in Yeast

The yeast \textit{Saccharomyces cerevisiae} contains an endogenous reductase capable of efficient electron transfer to P450s\textsuperscript{323} and was the first host where mammalian\textsuperscript{324,325} and plant\textsuperscript{326} P450s were successful expressed. \textit{S. cerevisiae} contains three P450s and has therefore also its own reductases. Additionally, yeast supports posttranslational machinery (e. g. glycosylation, membrane association, correct protein folding) for the expression of active eukaryotic P450s.
More than twenty years ago it was found that the P450 activity is limited by the quantity and availability of the P450 reductase. For example CYP71A1 from avocado, CYP73A1 from artichoke and CYP73A3 from alfalfa have been expressed successfully in *S. cerevisiae*, however activity was reduced. Possible reasons for this could be that the yeast reductase is inadequate for the amount of expressed P450, and that the electron transfer is not optimal between yeast reductase and plant P450. Thus the activity was increased when the partnering plant reductase was expressed simultaneously. The group of Ohkawa reported the successful expression of different mammalian membrane bound P450 together with the corresponding reductase in yeast. For example, the 17α-hydroxylase (P450c17) was ten times more active than the wild-type when the native P450c17 was fused to the 23 amino acid truncated yeast reductase. Plant P450 co-expression systems with its corresponding NADPH P450 reductase in yeast have been generated. Examples of co-expression include the CYP98A14 from *Coleus blumei*, which catalyses the 3-hydroxylation of 4-coumaroyl-3’,4’-dihydroxyphenyllactate and the 3’-hydroxylation of caffeoyl-4’-hydroxyl-phenyl lactate by forming rosmarinic acid.

One of the first vectors used for expression of plant P450s in yeast was the shuttle vector pYeDP60 (Figure 2.3) which has a high copy number when expressed in *E. coli*. This vector, which contains the galactose-inducible GAL10-CYC1 promoter is coupled with the yeast P450 reductase and is still used due to its high stability in rich medium at high cell density. To overcome coupling problems between the plant P450 and the yeast reductase, special yeast strains WAT11 and WAT21, which express the Arabidopsis NADPH P450 reductase ATR1 and ATR2, respectively, were constructed. Over twenty Arabidopsis P450s (CYP71B15, 71A1, 72B1, 73A5, 77A4, 77A6, 82C2, 82C4, 82F1, 82G1, 83A1, 83B1, 84A1, 85A1-2, 86A2, 86A8, 88A4, 94B1-3, 94C1, 86A1-4, 97B3, 98A3, 703A2, 704A2, 704B1, 707A1-4, 709C1, 711A1 and 734A1) have been successfully expressed using the pYeDP60 vector and the WAT11 strain.
The alternate *Pichia pastoris* system can also be used to overcome the expression problems described above for plant P450s in *E. coli*.\textsuperscript{342} The Arabidopsis CYP85A2 – which converts teasterone and also typhasterol to 7-oxa-teasterone and 7-oxatyphasterol, respectively – and ATR1 have been expressed simultaneously in *Pichia pastoris* and activity was found.\textsuperscript{343}

### 1.6 Aim of the project

The plant cytochrome P450 superfamily is a very attractive target for industrial applications, due to the high number of different reactions which they are able to catalyse. Cytochromes P450 work most efficiently when combined with their endogenous cytochrome P450 reductases as electron donor. The development of a fusion protein containing plant P450s and appropriate reductase to produce robust and efficient biocatalysts would be of relevance to industry.

The aim of this project is to develop a platform technology by cloning plant P450s for recombinant expression, for physiological characterisation studies and industrial application. Different expression systems were tested to get functional enzymes. The simplest expression system used was *E. coli*; however, as plant P450s are membrane bound, functional expression is more likely to occur in eukaryotic expression systems. Therefore, the yeast *S. cerevisiae* was also tested as host.
2 Materials and Methods

2.1 Reagents and consumables

The reagents and consumables used in this study were purchased from AccuStandard (New Haven, US), Bio-Rad Laboratories Ltd. (Hemel Hempstead, UK), Bruker Ldt. (Coventry, UK), Carl Zeiss Ltd. (Rugby, UK), Daicel Chemical Industries Ldt. (Tokyo, Japan), Difco (BD, Oxford, UK), Expedeon Ltd. (Harston, UK), Fisher Scientific UK Ltd. (Loughborough, UK), Finnzymes (Vantaa, Finland), Formedium (Norfolk, UK), Invitrogen Ltd. (Paisley, UK), Kartell spa (Noviglio, Italy), Levington F2 compost (Scotts, Bamford, Ipswich, UK), Matrix Science Ldt. (London, UK), Melford Laboratories Ltd. (Ipswich, Suffolk UK), Merck Chemicals Ltd. (Nottingham, UK), Millipore (Tullagreen, Ireland), New England Biolabs (Ipswich, UK), Oxoid Ltd. (Basingstoke, UK), Promega UK Ltd. (Southampton, UK), Qiagen Ltd. (West Sussex, UK), Sarstedt Ltd. (Leicester, UK), Sartorius AG (Göttingen, Germany), Scientific Laboratory Supplies Ltd. (Nottingham, UK), SelectScience Ltd. (Bath, UK), Sigma-Aldrich Company Ltd. (Dorset, UK), Starlab Ltd. (Milton Keynes, UK), Sterilin Ltd. (Newport, UK), Thermo Fisher Scientific p/a Perbio Science UK, Ltd. (Cramlington, UK), UVItec Ltd. (Cambridge, UK) and VWR International Ldt. (Lutterworth, UK).

TNT was donated by the Defence Science and Technology Laboratory (Dstl; Fort Halsted, UK), dinitrotoluences (Sigma-Aldrich), aminodinitrotoluenes by Supelco (Sigma-Aldrich) and hydroxyaminodinitrotoluenes by AccuStandard.

DNA Polymerases and restriction enzymes were bought from New England Biolabs and Finnzymes.

PCR Primers were synthesised by Sigma-Aldrich and genes by GeneArt (Regenburg, Germany).

The construct of the artificial fusion construct IFS-CPR (IFS = isoflavone synthase 1 from Glycine max fused to the P450 reductase from Catharanthus roseus) was kindly provided by Prof. Mattheos Koffas (State University of New York, Buffalo, USA) and the nicotine-N-demethylase (CYP82E4) from Nicotiana tabacum from Prof. Ralph Dewey (Department of Crop Science, North Carolina State University, Raleigh, USA).
The water used to produce buffers and solutions was molecular biology grade water purified using an Elga Purelab Ultra water polisher (Elga Labwater, High Wycombe, UK).

2.2 Media, strains and plasmids

2.2.1 Bacterial media

*Escherichia coli* was grown in Lysogeny broth (LB), Terrific broth (TB), auto-induction and M9 minimal media. Super Optimal broth with Catabolite repression (SOC) was used for recovering *E. coli* cells after plasmid DNA transformation.

**LB** contained 10 g/l tryptone (Formedium), 10 g/l sodium chloride (NaCl) (Fisher Scientific) and 5 g/l yeast extract (Formedium). For the preparation of the solid agar plates, 15 g/l agar (Formedium) was added to LB prior autoclaving and then poured into petri dishes (Sterilin).

**TB** contained 12 g tryptone (Oxoid), 24 g yeast extract (Oxoid), 4 ml Glycerol (Fisher Scientific) and water to a final volume of 900 ml and sterilised. Before use, 100 ml of sterile 10x TB salt (170 mM KH$_2$PO$_4$ and 720 mM K$_2$HPO$_4$, Fisher Scientific) were added.

**Auto-induction** medium$^{344}$ (1 l) was made of 950 ml ZY solution (10 g tryptone, 5 g yeast extract and water until a final volume of 950 ml), 25 ml 50x M-solution (0.25 M sodium sulphate (Fisher Scientific), 2.5 M ammonium chloride (Fisher Scientific), 1.25 M monopotassium phosphate (Fisher Scientific), 1.25 M disodium phosphate (Fisher Scientific)), 10 ml 50x 5052 solution (25% glycerol (Fisher Scientific), 0.14 M glucose (Sigma-Aldrich), 0.3 M α-Lactose (Sigma-Aldrich)), 1 ml 1 M MgSO$_4$ and 1 ml 1000x trace metal solution.

100 ml 1000x trace metal solution contained 0.25 M iron(III)chloride hexahydrate (Fisons) (in 0.1 M HCl), 20 mM calcium chloride (Fisher Scientific), 10 mM manganese(II)chloride tetrahydrate (Sigma-Aldrich), 10 mM zinc(II)sulphate (Fisher Scientific), 2 mM cobalt(II)chloride hexahydrate (Sigma-Aldrich), 2 mM copper(II)chloride dehydrate (Sigma-Aldrich), 2 mM nickel(II)chloride hexahydrate (Sigma-Aldrich), 2 mM sodium molybdate pentahydrate (Sigma-
Aldrich), 2 mM sodium selenite pentahydrate (Fluka, Sigma-Aldrich) and 2 mM boric acid (Fisher Scientific).

Five times M9 was made of 33.9 g/l Na$_2$HPO$_4$ (Fisher Scientific), 15 g/l KH$_2$PO$_4$, 2.5 g/l NaCl (Fisher Scientific), 5 g/l NH$_4$Cl (Fisher Scientific) and sterilised by autoclaving before use. For 1x M9 media was used 200 ml/l 5x M9, 800 ml/l water, 10 ml/l 1 M glucose (sterile filtrated, Fisher Scientific), 2 ml/l 1 M MgSO$_4$ (Fisher Scientific, sterile filtrated in Millex-MP filter unit, 0.22 μm from Millipore) and 100 μl 1 M CaCl$_2$ (sterile filtrated, Fisher Scientific).

SOC, used for transformation reactions, contained 5 g yeast extract (Formedium), 0.5 g NaCl (Fisher Scientific), 20 g tryptone (Formedium) and 950 ml water. After autoclaving, sterile filtrated glucose was added to a final concentration of 20 mM.

### 2.2.2 Bacterial strains

Different *E. coli* strains (Table 2.1) stored in 20% glycerol (Fisher Scientific) at -80 °C or after fresh transformation were streaked onto LB agar plates containing the appropriate antibiotics as necessary. Plates were incubated over night at 37 °C. One colony was used to inoculate a 10 ml starter culture, which was grown shaking at 250 rpm over night at 37 °C (Incubators used were from Gallenkamp, Jencons and Heraeus). Starter cultures were either used for plasmid preparation (Chapter 2.5, page 60) or to inoculate growth media for protein expression as described in Chapter 2.10.1 (page 67). TOP10, DH5α and NovaBlue were used for cloning and Rosetta 2 (DE3) and Rosetta gami 2 (DE3) for protein expression (Table 2.1).
# Chapter 2 – Materials and Methods

Table 2.1: List of E. coli strains used for cloning and recombinant expression in this project

<table>
<thead>
<tr>
<th>Strain name</th>
<th>Genotype</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOP10</td>
<td>F- mcrA Δ(mrr-hsdRMS mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 supG</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>DH5α</td>
<td>F- Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK-, mK+) phoA supE44 λ− thi-1 gyrA96 relA1</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>JM109</td>
<td>F’ [traD36 proA’B’ lacIΔ(lacZ)M15] supE44 endA1 hsdR17 gyrA96 relA1 thi Δ(lac-proAB)</td>
<td>Stratagene</td>
</tr>
<tr>
<td>BL21 (DE3)</td>
<td>F’ ompT hsdSB(rB-, mB+) gal dcm (DE3)</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>NovaBlue (DE3)</td>
<td>endA1 hsdR17 (rK12 mK12+) supE44 thi-1 recA1 gyrA96 relA1 lac F[proA’B’ lacIΔ(lacZ)M15::Tn10] (TetR)</td>
<td>Novagen-Merck</td>
</tr>
<tr>
<td>Rosetta 2 (DE3)</td>
<td>F’ ompT hsdS8a(rB- mB-) gal dcm (DE3) pRARE2 (CamR)</td>
<td>Novagen-Merck</td>
</tr>
<tr>
<td>Rosetta gami 2 (DE3)</td>
<td>Δ(ara–leu)7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL(DE3) F’[lacI ΔlacI pro] gor522::Tn10 trxB pRARE2 (CamR, StrR, TetR)</td>
<td>Novagen-Merck</td>
</tr>
</tbody>
</table>

The pRARE2 plasmid contains tRNA genes for the rare codons AUA (Ile), AGG (Arg), AGA (Arg), CUA (Leu), CCC (Pro), CGG (Arg) and GGA (Gly).
## 2.2.3 Plasmids for gene cloning and enzyme expression

Plasmids for gene cloning and enzyme expression are shown in Table 2.2.

<table>
<thead>
<tr>
<th>vector</th>
<th>antibiotics resistance</th>
<th>features</th>
<th>source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCR2.1-TOPO</td>
<td>Kanamycin</td>
<td>T-overhang vector for cloning PCR products, topoisomerase</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pTrHis2/lacZ</td>
<td>Ampicillin</td>
<td>contained <em>isoflavone synthase</em> 1 gene from <em>Glycine max</em>, vector pTrHis2/lacZ, original from Invitrogen</td>
<td>Prof. Koffas</td>
</tr>
<tr>
<td>LIC</td>
<td>Kanamycin</td>
<td>pETYSBLIC3C vector, cleavable his tagged N-terminus</td>
<td>Dr. Gideon Grogan</td>
</tr>
<tr>
<td>LicRed</td>
<td>Kanamycin</td>
<td>pETYSBLIC3C vector, cleavable his tagged N-terminus, contains RhF reductases from <em>Rhodococcus</em> sp.</td>
<td>Dr. Federico Sabbadin</td>
</tr>
<tr>
<td>lamATR2tr</td>
<td>Kanamycin</td>
<td>pETYSBLIC3C vector, cleavable his tagged N-terminus, contains lambda linker and truncated ATR2 reductases from Arabidopsis</td>
<td>this work, chapter 5</td>
</tr>
<tr>
<td>licATR2tr</td>
<td>Kanamycin</td>
<td>pETYSBLIC3C vector, cleavable his tagged N-terminus, contains lic linker and truncated ATR2 reductases from Arabidopsis</td>
<td>this work, chapter 5</td>
</tr>
<tr>
<td>pYeDP60</td>
<td>Ampicillin</td>
<td>GAL10-CYC1 promoter</td>
<td>Prof. Daniele Werck-Reichart</td>
</tr>
</tbody>
</table>

Table 2.2: Plasmids for gene cloning and enzyme expression
The pTrcHis2/lacZ-vector (Figure 2.1, Invitrogen) containing the fusion IFS-CPR was derived from Prof. Mattheos Koffas (State University of New York, Buffalo, USA).306

Figure 2.1: Vector map for pTrcHis2/lacZ
(pTrc = trc promoter, oLAC = Lac operator, RBS = ribosome binding site, ATG = expression start codon, MCS = multible cloning site, myc = myc epitope, His = 6x histidines residues, lacZ = β-galactosidase gene, term = terminator, ampR = ampicillin resistance gene, pBR322 ori = origin of replication of the plasmid pBR322, LacI = repressor gene for IPTG induction)
The LIC-vector (Figure 2.2), LicRed\textsuperscript{146} and the in this work developed lamATR2tr and licATR2tr vector are based on pETYSBLIC3C vector\textsuperscript{345}, which contains an additional cleavable his tagged N-terminus.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{ LIC-vector.png}
\caption{Vector map of LIC-vector \hfill
\textit{(pT7 = T7 promotor, oLAC = Lac operator, RBS = ribosome binding site, His = 6x histidines residues, which can be cleaved HRV 2C protease site = 3C, tT7 = T7 terminator, f1 ori = f1 phage origin of replication, pBR322 ori = origin of replication of the plasmid pBR322, kanR = kanamycin resistance gene, LacI = repressor gene for IPTG induction)}}
\end{figure}
The vector pYeDP60 (vector map Figure 2.3) is a shuttle vector for cloning in *E. coli* and an expression in yeast.  

**Figure 2.3: pYeDP60 shuttle vector map with an origin of replication for *E. coli* and yeast**  
(tPGK = phosphoglycerate kinase terminator, pGAL = galactose promoter, ADE 2d and URA3 = selection marker for adenine and uracil auxotrophy, ampR = ampillin resistance gene)

### 2.3 Preparation of chemically competent *Escherichia coli* cells

All steps were performed with the appropriate antibiotics for the different *E. coli* strains (Table 2.1) in the growth media. For the preparation of chemically competent cells, *E. coli* strains were streaked on agar plates and incubated overnight at 37 °C. A starter culture (10 ml LB inoculated by a single colony from plate) was used to inoculate 100 ml LB to an OD$_{600}$ of 0.1. This culture was grown to an OD$_{600}$ of 0.4-0.5 and then centrifuged in 25 ml batches in 50 ml falcon tubes (Sarstedt) at 4000 rpm for 5 min and 4 °C in a CR312 Swinging Bucket Rotor Centrifuge (Jouan). The pellets were resuspended in 10 ml of an ice cold, sterile 50 mM MgCl$_2$/20 mM CaCl$_2$ solution per tube and incubated on ice for 30 min. Cells were then centrifuged at 4000 rpm for 5 min (CR312 Swinging Bucket Rotor Centrifuge, Jouan) and the pellet resuspended in 1 ml ice cold,
sterile 100 mM CaCl₂ (Fisher Scientific). Sterile glycerol (10% v/v Fisher Scientific) was added and the cells were incubated on ice for further 30 min. Aliquots of 50 µl were snap frozen in liquid nitrogen and stored at -80 ºC until use.

2.4 Plant media, growth conditions and strains

Wild type *Arabidopsis thaliana* (Columbia 0, NASC Stock acc. No. N50193) was used for the isolation of CYP81D8, CYP81D11 and the two Arabidopsis reductases ATR1 and ATR2. All steps were carried out under sterile conditions by preparing media and handling plant material in the laminar flow cabinet.

Arabidopsis was grown in half MS (0.215% Murashige and Skoog Basal Salt mixture, Sigma-Aldrich) liquid culture containing sucrose (0.68%, Fisher Scientific), adjusted to pH 5.7 with KOH (Fisher Scientific) and if required 60 µM TNT. To solidify the medium, 0.8% (w/v) agar (Formedium) was added.

Seeds were sterilised in a fume cupboard as follows: Seeds were transferred to 1.5 ml micro tubes (Sarstedt) which were placed with open lids into a plastic box. The lid of the plastic box was closed immediately after the addition of 3 ml concentrated hydrochloric acid to 100 ml Bleach haychlor (Scientific Laboratories Supplies) in a glass beaker. After 3 h the lid of the plastic box was opened, the micro tubes were closed and then placed into the sterile flow hood with opened lids for 10 min to allow evaporation of remaining hydrogen chloride. Seeds were then transferred to half MS agar plates and imbibed at 4 ºC for at least three nights.

Eight one-day-old seedlings were transferred from agar plates into 100 ml sterile glass flasks containing 10 ml half MS. Seeds were grown for a further 13 days under light condition of 80 µmol·m⁻²·s⁻¹, 16 h photoperiod at 25 ºC and 130 rpm.
2.5 Plasmid DNA preparation

For a plasmid preparation, 5 ml cultures of *E. coli* were grown overnight at 37 °C, 200 rpm, in LB containing the appropriate antibiotics. The cultures were harvested by centrifugation at 4000 rpm for 10 min (Jouan Model CR 312 Centrifuge, Swinging Bucket Rotor, SelectScience). The plasmids were purified using the QIAprep Spin Miniprep kit (Qiagen) according to the manufacturer’s instructions.

2.6 Purification of DNA fragments

The DNA fragments resulting from PCR amplification and restriction digest were purified using the Wizard SV Gel and PCR Clean-Up System (Promega), according to the manufacturer’s protocol. Where required, DNA fragments were separated using agarose gel electrophoresis and excised from the gel before purification.

2.7 Detection of the mRNA transcript

2.7.1 RNA extraction from *Escherichia coli*

The RNA was extracted by using the RNeasy Mini Kit from Qiagen for RNA isolation following by a reverse transcription PCR reaction.

For the RNA isolation, *E. coli* Rosetta 2 (DE3) cells containing the plasmid with the gene of interest were grown in 100 ml M9 media (37 °C, 200 rpm) until the O\textsubscript{D}\text{600} reached 0.6-0.8 and then induced by adding 1 mM IPTG, 1 mM δ-aminolevulinic acid (ALA), 0.5 mM FeCl\textsubscript{3} (Fisons), 5 μg/l riboflavin and incubated at 20 °C shaking (200 rpm) overnight.

The RNA was protected by using 2 ml RNAprotect Bacteria Reagent (Qiagen) to 1 ml culture, vortexed for 5 sec, incubated for 5 min at room temperature and centrifuged at 5000x g for 10 min (bench top microlitre centrifuge Sigma 1-15P). The supernatant was discarded and 200 μl TE buffer (30 mM Tris-HCl, 1 mM EDTA, pH 8.0) containing 0.4 mg proteinase K and 4 mg lysozyme were added to the cell pellet, resuspended by pipetting and incubated for 10 min at room temperature (vortex at least every two min). Then, 700 μl RTL buffer containing
β-mercaptoethanol (10 μl β-mercaptoethanol/1 ml RTL) were added, vortexed, centrifuged for 5 min at full speed (bench top microlitre centrifuge Sigma 1-15P) and the supernatant transferred into a new tube. After adding 500 μl 100% ethanol and mixing by inverting, the solution was transferred to a column, spun at 8000x g for 15 sec (bench top microlitre centrifuge Sigma 1-15P) and the flow-through discarded. The column was washed using 350 μl RW1 buffer and spun 15 sec at 8000x g. DNA was removed by adding 70 μl RDD buffer containing 10 μl DNAse stock (RNase free DNase Set, Qiagen) and incubated for 15 min at room temperature. The column was washed twice as described above with 350 μl RW1 buffer followed by adding 500 μl RW1 buffer to the column. The column was transferred to a 1.5 ml collection tube, 45 μl RNase free water was added, spun 1 min at 8000x g and the RNA concentration measured with the Nanodrop ND-1000 spectrophotometer.

2.7.2 RNA extraction from plant tissue
Fourteen-day-old liquid culture grown Arabidopsis plants were harvested after six hours incubation with 60 μM TNT (in 10 μl DMSO). 10 μl of DMSO alone were added to the cultures for negative control. Arabidopsis RNA was isolated by grinding eight whole plants in a pestle and mortar cooled in liquid nitrogen. The RNA was isolated using a plant RNAeasy Mini Kit from Qiagen according to the manufacturer’s instructions. An on-column, DNA digestion using RNAase-free DNAase (Qiagen) was included. The RNA content was quantified using a Nanodrop ND-1000 Spectrophotometer.

Ten μg RNA were added to a total of 24 μl RNase free water. 1 μl of oligo dT primer (Invitrogen) was added and the mixture was incubated in the thermocycler for 2 min at 95 °C and then chilled on ice. Then the following substances were added: 8 μl 5 x 1st strand buffer (Invitrogen), 2 μl 2.5 mM dNTPs (Bioline), 2 μl 0.1M DTT (Invitrogen), 1 μl RNAsin (RNAse inhibitor), 2 μl Superscript II (Invitrogen). The total volume was brought to 40 μl using RNase free water. The mixture was incubated at 42 °C for 2 h. Samples were frozen at -80°C.
2.7.3 Reverse transcription from plant mRNA to cDNA

To 10 μg plant mRNA was added 1 μl oligo dT primer (0.5 μg/μl, Invitrogen) and RNase free water to a final volume of 25 μl, incubated at 95 ºC for 2 min and then chilled on ice. After adding 8 μl 5x 1st strand buffer (Invitrogen), 125 μM dNTPs (Invitrogen), 5 μM DTT (Invitrogen), 40 U RNAsin ribonuclease inhibitor (Promega), 400 U Superscript II (Invitrogen), the total volume was brought to 40 μl using RNase free water and the solution was incubated at 42 ºC for 2 h. Samples were frozen at -80 ºC until use for PCR.

2.8 Traditional cloning method

2.8.1 Designing primers for truncated P450 and reductase genes

Plant P450s and P450 reductases are membrane associated through an N-terminal hydrophobic membrane anchor of around 25 to 70 amino acids. The association to the membrane is causing mainly the insolubility of these proteins. The catalytic centre is situated in the cytoplasm and a removal of the N-terminus potentially increasing the solubility will not reduce the enzyme activity. To design forward primers, the amino acid sequence was analysed with TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) and SignalP3.0 software (http://www.cbs.dtu.dk/services/SignalP/). The TMHMM program predicts membrane regions and SignalP 3.0 can identify the location of a signal peptide and cleavage sites.

2.8.2 Polymerase chain reaction (PCR)

Primers were designed using primer3 software (http://frodo.wi.mit.edu) and synthesised by Sigma-Aldrich.

PCRs for amplifying the genes of interest were performed using the Phusion DNA Polymerase (Finnzymes), which is a highly accurate proof reading polymerase in a Peltier Thermal Cycler PTC-200. PCR amplifications were performed in volumes ranging from 25 μl to 50 μl using 5x buffer HF. Primers were employed at an ultimate concentration of 400 nM, dNTPs (Invitrogen) at a final concentration of 200 μM each and template DNA at 200 ng.
The programme used was, if not otherwise indicated, 5 min at 98 °C for denaturation, 30 cycles of 10 sec denaturing at 98 °C, 30 sec annealing at 60-65 °C, 1.00-1.20 min extension at 72 °C and 10 min at 72 °C. PCR products were stored at -20 °C and a part of each PCR product was visualised on agarose gels. By appearance of unspecific PCR products, 400 mM NDSB 201 (3-(1-Pyridinio)-1-propanesulfonate) was added to improve the specificity of the PCR reaction.

The Taq polymerase (New England Biolabs) was used for colony PCR after a transformation reactions to identify the presence of the gene. Therefore, 2 μl cell template from a single colony (preparation according to Chapter 2.8.5, page 64) were mixed with 2.5 μl 10x Thermo pol buffer, 200 μM dNTP mix (Invitrogen), 44 mM DMSO, 400 nM primer (forward and reverse) in a total volume of 25 μl. The PCR used a 5 min denaturation at 94 °C, followed by 35 cycles of 94 °C denaturation for 15 sec, 55 °C annealing for 30 sec and 72 °C extension for 1.00-1.20 min and a final extension step of 72 °C for 10 min.

2.8.3 Preparation of PCR products for cloning

Immediately after purifying the PCR product (see Chapter 2.6, page 60), A-tail overhangs were added to create blunt-ended PCR products. 3 μl cleaned PCR product (negative control with water) was added to 0.5 μl 10 x thermo pol buffer, 0.2 mM dATPs, 0.25 U Taq-polymerase and water to a final volume of 5 μl. After incubation at 72 °C for 10 min, 1 μl TOPO vector and 1 μl salt solution (both from TOPO TA Cloning Kit from Invitrogen) was added and the whole mix placed at room temperature for 30 min.

2.8.4 Transformation of plasmid DNA into chemically competent

*Escherichia coli* cells

An aliquot 50 μl of frozen competent *E. coli* DH5α was defrosted on ice for 10-15 min and then 2 μl of plasmid DNA was added, mixed by stirring with pipette and left on ice for 5 min. *E. coli* cells were then heat-shocked, in a water bath at 42 °C for 40 sec and again chilled on ice. After 5 min, 200 μl SOC-medium was added and the freshly transformed cells were allowed to recover at 37 °C under shaking. After 1 h, aliquots were spread onto LB agar plates, containing antibiotics if necessary and grown at 37 °C overnight. For TOPO blue / white
screening cells were spread onto plates containing 64 μg/ml 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal).

2.8.5 Screening transformants for recombinant genes by PCR

Single colonies were picked with a pipette tip and resuspended in 20 μl sterile water. Bacteria cell walls were destroyed by incubation at 95 °C for 5 min and then cycled 5 times through 15 °C for 30 sec followed by 95 °C for 30 sec and then frozen at -80 °C for 10 min. 2 μl of this solution was used as template for the PCR reaction using Taq polymerase (Chapter 2.8.2, page 62).

2.8.6 Restriction endonuclease digest of DNA

Restriction enzymes were obtained from New England Biolabs and used with the buffer system supplied by the manufacturer. Digests were used to create diagnostic restriction maps for construct analysis or in preparation for subsequent cloning steps. For restriction mapping 200 ng DNA was digested with 10-20 U of enzyme in a 10 μl reaction and if necessary 1 μg BSA was added. For preparation of DNA fragments 1-5 μg DNA was digested with 5-10 U of enzyme and 5 μg BSA. All reactions were incubated for 1.5 h at 37 °C. For restriction mapping, 5 μl of the reaction were analysed using agarose gel electrophoresis (0.8% agarose gel, see Chapter 2.11, page 72). For preparative digests, products of the reaction were purified, if necessary using gel purification (Chapter 2.6, page 60). When buffering systems for digests using multiple enzymes, which were not compatible, sequential digests with both enzymes on the same plasmid were conducted.

2.8.7 Dephosphorylation of DNA 5’ end of the vector

The dephosphorylation of the DNA 5’ ends was performed using 45 μl digested vector (cleaned from gel after digestion), 5 μl 10x arctic phosphatase buffer and 1 μl arctic phosphatase (5000 U/ml, New England Biolabs). The samples were incubated at 37 °C for 45 min and then the phosphatase was heat inactivated at 65 °C for 5 min before the solution was used for the ligation reaction (Chapter 2.8.8).
2.8.8 DNA ligation reaction

Ligation reactions were performed with dephosphorylated plasmid DNA (Chapter 2.8.7) and phosphorylated insert DNA. Ligation was performed on 100 ng of DNA, in 3:1, 1:1 or 1:3 ratios of vector to insert. Controls included phosphorylated and dephosphorylated vector only. Ligation reactions were performed using T4 DNA ligase (New England Biolabs). The reactions were buffered in 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP (Invitrogen), 10 mM dithiothreitol (DTT, Invitrogen), pH 7.5 at 25 °C and 400 U of ligase used in a total volume of 20 µl.

2.8.9 DNA sequencing and analysis

Sequencing reactions were performed by the Genomics Laboratory, Technology Facility, University of York (York, UK). A list of primers used for sequencing is shown in Table 2.3. PCR primer of the insert were used for the sequencing of the yeast shuttle vector pYeDP60. Sequence chromatograms were analysed using Applied Biosystems Sequence Scanner 1.0 and sequences were aligned using the BLAST packages available at the National Centre for Biotechnology information (www.ncbi.nlm.nih.gov/blast/).

<table>
<thead>
<tr>
<th>name</th>
<th>sequence (5'--&gt;3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>M13 forward</td>
<td>GAAAAACGACGGCCAGTG</td>
</tr>
<tr>
<td>M13 reverse</td>
<td>GGAAACAGCTATGACCATG</td>
</tr>
<tr>
<td>T7</td>
<td>TTATACGACTCACTATAGGG</td>
</tr>
<tr>
<td>T7term</td>
<td>TATGCTAGTTATTGCTAGCGGT</td>
</tr>
<tr>
<td>ATR2seq R</td>
<td>CATAAACGTGCTTTTGACTTCACCA</td>
</tr>
<tr>
<td>RhF rev</td>
<td>TGCCGCCCGCGACGAACGTAGTGCTC</td>
</tr>
</tbody>
</table>
2.9 Ligation independent cloning method

The Ligation Independent Cloning (LIC) enables a rapid high-throughput cloning without any ligation steps due to the complementary long overhangs on insert and vector (Figure 2.4).\textsuperscript{345}

![Figure 2.4: Schematic procedure for ligation independent cloning](image)

Therefore, the LIC-vector, based on pETYSBLIC3C vector\textsuperscript{345} containing six cleavable N-terminal histidines, was digested using BseR1. Briefly, 100 ng LIC-vector, 10 μl NEB4 buffer, 20 U BseRI and water to a final volume of 100 μl were incubated for 1 h 50 min, then separated on a 0.6% agarose gel and purified (Chapter 2.6, page 60). The insert was amplified by PCR using Phusion polymerase (Chapter 2.8.2, page 62) and primers designed using HiTel software (TF Protein Production Laboratory, University of York, http://biolfwsi.york.ac.uk/cgi-bin/primers.cgi?). The purified insert and vector were treated with T4 polymerase (LIC qualified, Novagen-Merck) and the complementary bases A and T (Table 2.4) to create the long sticky ends (Figure 2.5). The reaction was performed at 22 °C for 30 min and stopped with further incubation of 75 °C for 20 min.
### Table 2.4: Protocol for T4 polymerase treatment

<table>
<thead>
<tr>
<th>vector</th>
<th>insert</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 pmol</td>
<td>linearised vector</td>
</tr>
<tr>
<td>40 μl</td>
<td>10x T4 pol buffer</td>
</tr>
<tr>
<td>20 μl</td>
<td>100 mM DTT</td>
</tr>
<tr>
<td>10 μl</td>
<td>100 mM dTTP</td>
</tr>
<tr>
<td>8 μl</td>
<td>T4 DNA polymerase*</td>
</tr>
<tr>
<td>water</td>
<td>to a final volume of 400 μl</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2 pmol</td>
<td>insert</td>
</tr>
<tr>
<td>2 μl</td>
<td>10x T4 pol buffer</td>
</tr>
<tr>
<td>1 μl</td>
<td>100 mM DTT</td>
</tr>
<tr>
<td>0.5 μl</td>
<td>100 mM dATP</td>
</tr>
<tr>
<td>0.4 μl</td>
<td>T4 DNA polymerase*</td>
</tr>
<tr>
<td>water</td>
<td>to a final volume of 20 μl</td>
</tr>
</tbody>
</table>

* 2.5U/μl LIC qualified T4 DNA polymerase, Novagen/Merck

---

**Figure 2.5: LIC-vector after T4 /dTTP treatment**
(BseRI restriction site in italics)

The T4 treated vector was purified (Chapter 2.6, page 60) and diluted to a concentration of 50 ng/μl. For the ligation, 1 μl vector and 2 μl insert (negative control: 2 μl water) were mixed and incubated at room temperature for 10 min. To enable easy transformation, divalent ions such as Mg$^{2+}$ were complexed by adding 1 μl 25 mM EDTA to the mix and incubated 10 min on ice. 2 μl of ligation reaction was added to 50 μl chemical competent *E. coli* DH5α cells (Chapter 2.8.4, page 63), incubated for 5 min on ice, heat shocked for 40 sec at 42 °C and chilled for 5 min on ice. SOC medium (200 μl) was added and incubated shaking (200 rpm) for 1 h at 37 °C. The DNA will be ligated by repair enzymes, which are naturally present in *E. coli* cells. The sample was spread onto agar plates containing 50 μg/ml kanamycin and incubated overnight at 37 °C. Positive transformants were analysed using the protocol transcribed in Chapter 2.8.5, 2.8.6 and 2.8.9.

### 2.10 Protein expression and purification

#### 2.10.1 Expression in *Escherichia coli*

The strain *E. coli* Rosetta 2 (DE3) (Table 2.1) was transformed with the plasmid containing the gene of interest (Chapter 2.8.4, page 63). Positive transformants were grown in 10 ml LB containing 34 μg/ml chloramphenicol and 100 μg/ml
kanamycin after confirmation of the presence of the insert by PCR and incubated overnight at 37 °C and 200 rpm. M9 minimal medium was used to grow bacteria cells expressing P450s and fusion enzymes and LB medium for ATR1tr or ATR2tr. The main culture was inoculated with the preculture at an OD$_{600}$ of 0.1 and incubated shaking at 37 °C until an OD$_{600}$ = 0.6-0.8. Protein expression was then induced by adding 1 mM IPTG. Additionally, 5 μg/l riboflavin was added for reductases as precursor of FMN and FAD and 0.5 mM FeCl$_3$ (Fisons), 1 mM ALA (δ-aminolevulinic acid, a heme ring precursor) for P450 expression. The proteins were expressed shaking (200 rpm) at 15 °C or 20 °C overnight.

2.10.2 Solubilisation buffers

To increase the solubility of the truncated Arabidopsis reductases, different solubilisation buffers were tested following the protocol from Lindwall et al. 2000$^{347}$. The ATR1tr was transformed into E. coli Rosetta 2 (DE3) and expressed in LB medium under the conditions described in Chapter 2.8.4. Cells were harvested and resuspended in 35 ml 10 mM Tris, pH 8.5 containing 100 mM NaCl and 1 mM EDTA and then divided in 30 1 ml aliquots. Cells were spun (2300x g, 1 min), the supernatant discarded and the cell pellets resuspended in 1 ml of the different solubilisation buffers listed in Table 2.5. Buffer 0 was the buffer used for the batch purification of the Arabidopsis reductases in Chapter 2.10.4, which resulted in precipitated protein after purification.
<table>
<thead>
<tr>
<th>buffer</th>
<th>Reagents to solubilise overexpressed protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>buffer 0</td>
<td>HIS-binding buffer: 50 mM sodium phosphate pH 8.0, 300 mM NaCl</td>
</tr>
<tr>
<td>buffer 1</td>
<td>100 mM Tris, 10 % glycerol, pH 7.6</td>
</tr>
<tr>
<td>buffer 2</td>
<td>100 mM Tris, 50 mM LiCl, pH 7.6</td>
</tr>
<tr>
<td>buffer 3</td>
<td>100 mM HEPES, 50 mM (NH₄)₂SO₄, 10 % glycerol, pH 7.0</td>
</tr>
<tr>
<td>buffer 4</td>
<td>100 mM HEPES, 100 mM KCl, pH 7.0</td>
</tr>
<tr>
<td>buffer 5</td>
<td>100 mM Tris, 50 mM NaCl, 10 % isopropanol, pH 8.2</td>
</tr>
<tr>
<td>buffer 6</td>
<td>100 mM K₂HPO₄/KH₂PO₄, 50 mM (NH₄)₂SO₄, 1% Triton X-100, pH 6.0</td>
</tr>
<tr>
<td>buffer 7</td>
<td>100 mM triethanolamine, 100 mM KCl, 10 mM DTT, pH 8.5</td>
</tr>
<tr>
<td>buffer 8</td>
<td>100 mM Tris, 100 mM sodium glutamate, 10 mM DTT, pH 8.2</td>
</tr>
<tr>
<td>buffer 9</td>
<td>250 mM K₂HPO₄/KH₂PO₄, 0.1 % CHAPS, pH 6.0</td>
</tr>
<tr>
<td>buffer 10</td>
<td>100 mM triethanolamine, 50 mM LiCl, 5 mM EDTA, pH 8.5</td>
</tr>
<tr>
<td>buffer 11</td>
<td>100 mM sodium acetate, 100 mM glutamine, 10 mM DTT, pH 5.5</td>
</tr>
<tr>
<td>buffer 12</td>
<td>100 mM sodium acetate, 100 mM KCl, 0.1% n-octyl-β-D-glucoside, pH 5.5</td>
</tr>
<tr>
<td>buffer 13</td>
<td>100 mM HEPES, 1 M MgSO₄, pH 7.0</td>
</tr>
<tr>
<td>buffer 14</td>
<td>100 mM HEPES, 50 mM LiCl, 0.1% CHAPS, pH 7.0</td>
</tr>
<tr>
<td>buffer 15</td>
<td>100 mM K₂HPO₄/KH₂PO₄, 2.5 mM ZnCl₂, pH 4.3</td>
</tr>
<tr>
<td>buffer 16</td>
<td>100 mM Tris, 50 mM NaCl, 5 mM calcium acetate, pH 7.6</td>
</tr>
<tr>
<td>buffer 17</td>
<td>100 mM triethanolamine, 50 mM (NH₄)₂SO₄, 10 mM MgSO₄, pH 8.5</td>
</tr>
<tr>
<td>buffer 18</td>
<td>100 mM Tris, 100 mM KCl, 2 mM EDTA, 1% Triton X-100, pH 8.2</td>
</tr>
<tr>
<td>buffer 19</td>
<td>100 mM sodium acetate, 1 M MgSO₄, pH 5.5</td>
</tr>
<tr>
<td>buffer 20</td>
<td>100 mM Tris, 2 M NaCl, 0.1% n-octyl-β-D-glucoside, pH 7.6</td>
</tr>
<tr>
<td>buffer 21</td>
<td>100 mM Tris, 1 M (NH₄)₂SO₄, 10 mM DTT, pH 8.2</td>
</tr>
<tr>
<td>buffer 22</td>
<td>100 mM sodium acetate, 50 mM LiCl, 5 mM calcium acetate, pH 5.5</td>
</tr>
<tr>
<td>buffer 23</td>
<td>100 mM HEPES, 100 mM sodium glutamate, 5 mM DTT, pH 7.0</td>
</tr>
<tr>
<td>buffer 24</td>
<td>100 mM triethanolamine, 100 mM sodium glutamate, 0.02% n-octyl-β-D-glucoside, 10% glycerol, pH 8.5</td>
</tr>
<tr>
<td>buffer 25</td>
<td>100 mM Tris, 50 mM NaCl, 100 mM urea, pH 8.2</td>
</tr>
<tr>
<td>buffer 26</td>
<td>100 mM triethanolamine, 100 mM KCl, 0.05% dextran sulfate, pH 8.5</td>
</tr>
<tr>
<td>buffer 27</td>
<td>100 mM K₂HPO₄/KH₂PO₄, 50 mM (NH₄)₂SO₄, 0.05% dextran sulfate, pH 6.0</td>
</tr>
<tr>
<td>buffer 28</td>
<td>100 mM HEPES, 50 mM LiCl, 0.1% deoxycholate, pH 7.0</td>
</tr>
<tr>
<td>buffer 29</td>
<td>100 mM Tris, 100 mM KCl, 0.1% deoxycholate, 25% glycerol, pH 7.6</td>
</tr>
<tr>
<td>buffer 30</td>
<td>100 mM potassium acetate, 50 mM NaCl, 0.05% dextran sulfate, 0.1% CHAPS, pH 5.5</td>
</tr>
</tbody>
</table>
Lysozyme (from chicken egg white, Sigma) was added and the suspension incubated for 5 min on ice. Cells of each aliquot were lysed (Misonix S-4000 sonicator at 70% maximum amplitude, 1 sec sonication bursts, 4 sec cooling intervals at 0 °C for a total processing time of 1 min) and then incubated gently shaking for 10 min at 4 °C. After centrifugation at 16000×g for 10 min, the soluble proteins in the supernatant were analysed by Bradford assay (Chapter 2.12), SDS PAGE (Chapter 2.13) and western blot (Chapter 2.14).

2.10.3 Cell lysis by sonication

1 l cell culture was harvest by centrifugation (7 min at 2500 xg, High Speed Sorvall RC5B+rotor, SLC-1500) after 20 h of expression and resuspended in 30 ml HIS-binding buffer (HIS-Binding buffer: 50 mM sodium phosphate pH 8, 300 mM NaCl) or buffer 18A (100 mM Tris, 100 mM KCl and 1% Triton X-100, pH 8.2) containing 37.5 μl 0.1 M PMSF (protease inhibitor solved in isopropanol). Lysozyme (from chicken egg white, Sigma) was added to a final concentration of 2 μg/ml to the cells and incubated on ice for 5 min. Cells were lysed usually in 35 ml batches with a Misonix S-4000 sonicator at 70% maximum amplitude alternating 3 sec sonication bursts with 7 sec cooling intervals at 0 °C for a total processing time of 4 min. The effect of the sonication was visualised by light microscopy (Carl Zeiss, Axiovert 200, AxioCam HRm) (Figure 2.6).

![Figure 2.6: E. coli Rosetta 2 (DE3) cells A before and B after sonication](image)
2.10.4 Protein purification in a batch process

All steps to purify the His-tagged protein were performed at room temperature and individual samples were taken for analysis by SDS PAGE (Chapter 2.13) and western blot analysis (Chapter 2.14). Protein purification was performed in a batch process by using His-Select Nickel Affinity Gel (Sigma) in a 50 ml falcon tube (500 μl slurry per 10 ml supernatant containing the protein). The resin was equilibrated with 10 resin volumes of buffer, mixed by inverting, spun at 1200 xg for 1 min (Jouan Model CR 312 Centrifuge, Swinging Bucket Rotor) and the supernatant was discarded. This step was repeated twice. The soluble protein fraction after sonication (Chapter 2.10.3, page 70) containing the protein of interest was applied to the resin material and gently shaken at room temperature for 1 h up to 4 h. After centrifugation at 1200 xg (Jouan Model CR 312 Centrifuge, Swinging Bucket Rotor), the supernatant was removed and the resin washed three times with fifteen resin volumes of buffer. To remove non specific bound protein, the resin was washed with one resin volume of buffer containing 5 mM imidazole (Fisher Scientific) shaking for 5 min. After centrifugation (1200 xg, 5 min, Jouan Model CR 312 Centrifuge), the supernatant was discarded and the His-tagged protein eluted by incubation of one resin volume of buffer containing 500 mM imidazole (Fisher Scientific) for 10 min to 1 h by gently shaking. After centrifugation (2400 xg, 5 min, Jouan Model CR 312 Centrifuge, Swinging Bucket Rotor), the supernatant was transferred to a spin column (Agilent Technologies, 0.22 μm cellulose acetate) to remove all resin traces by spinning 2 min at 13000 rpm (table centrifuge). The protein solution was dialysed twice against 5 l potassium phosphate (50 mM, pH 7.0) at 4 °C, if not otherwise specified.

2.10.5 Protein purification in a continuous process

A column (Chromatography column from Novagen) containing 1.5 ml His-Select Nickel Affinity Gel (Sigma) was equilibrated with 10 ml buffer. The soluble supernatant containing the protein of interest was applied to the resin material and then washed twice with 10 ml buffer. Unspecific bound protein was eluted with 5 ml of 5 mM imidazole (Fisher Scientific) in buffer and the His-tagged protein with 500 mM imidazole in buffer. The eluted protein was dialysed
against two batches 5 l potassium phosphate (50 mM, pH 7.0) at 4 °C, if not otherwise specified.

2.11 Agarose gel electrophoresis

Agarose gel electrophoresis was used to visualise and to separate digestion and PCR products. The concentration of agarose (Melford) was varied between 0.8-1.0% (w/v), according to the size of the fragments to be separated. Agarose was dissolved in 1x Tris-acetate-EDTA (TAE) buffer, made from a 50x TAE stock, which contained: 242 g/l Tris (Invitrogen), 57.1 ml/l glacial acetic acid (Fisher Scientific) and 100 ml/l 500 mM EDTA (Sigma-Aldrich), pH 7.0. To visualise DNA under UV light (UVItec system and the software UVIpro) ethidium bromide (Fluka, Sigma-Aldrich) was added to a final concentration of 0.2 µg/ml. DNA was diluted in 6x loading dye comprising 0.25% w/v bromophenol blue and 30% glycerol. To determine the product size 1 kb DNA ladder or 2-Log DNA Ladder (both from New England Bioloabs) were used according to the manufacturer’s instruction. The DNA was typically separated using a voltage of 90-130 V.

2.12 Protein detection

Protein concentrations were determined using the Coomassie (Bradford) Protein Assay Kit (Thermo Fisher Scientific). For quantification a standard curve was prepared using bovine serum albumin (BSA) solutions of known concentrations between 0-2 mg/ml measured at 595 nm in the spectrophotometer Varian Cary 50 Bio UV/Vis Spectrophotometer (Agilent Technologies). Measurements were taken in triplicates.

2.13 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

2.13.1 SDS PAGE with the Bio-Rad system

Standard procedures were followed according to Laemmli. Gels were poured and run with the Mini-Protean 3 system (Bio-Rad). Gels were polymerised with
12% w/v acrylamide (Fisher Scientific) in the resolving gel and 4% w/v acrylamide in the stacking gel. Four times SDS loading buffer contained 2 ml of 625 mM Tris (Invitrogen), pH 6.8, 4 ml of 10% (w/v) SDS (Melford), 2 ml of glycerol (Fisher Scientific), 1 ml of β-mercaptoethanol (Sigma-Aldrich), 30 mg of bromphenol blue (Sigma-Aldrich), 1 ml of H₂O. Two times loading buffer was added in equivalent volumes to the samples, denatured for 5 min at 100 °C, cooled to room temperature, centrifuged to bring down the condensation and pellet any insoluble substances and 15 to 20 μl loaded on the gel. Running buffer contained 25 mM Tris, 180 mM glycine (Fisher Scientific), 0.1% (w/v) SDS (pH 8.3). Electrophoresis was performed for approximately 45 min at 200 V. Prestained protein marker was purchased from New England Biolabs or Fermentas and gels were stained with Instant Blue (Expedeon). An UVItec system and the software UVIpro was used for the documentation.

2.13.2 SDS PAGE with the RunBlue system from Expedeon

SDS gradient gels (4 to 20% (w/v) acrylamide) were purchased from Expedeon. As running buffer was used 1x RunBlue RAPID. Twenty times of RunBlue RAPID was made of 0.6 M MOPS (Sigma), 1.2 M Tris (Invitrogen), 2% (w/v) SDS (Melford) and 130 mM sodium bisulphate (Expedeon). The sample preparation was done following the manufacturer’s instructions. The gels were run at 90 mA and then stained with Instant Blue (Expedeon) and documentated using an UVItec system and the software UVIpro.

2.14 Western blot analysis

For the western blot transfer a Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad) and a Model 200/2.0 power supply (Bio-Rad) were used. All western Blot analysis steps were performed at room temperature under gently shaking on a gel rocker (Stuart, Bibby Scientific Limited). An UVItec system and the software UVIpro was used for the documentation of the western Blot results.

2.14.1 Protein detection with anti-poly Histidine peroxidase conjugate

Protein SDS gels and nitrocellulose membrane (0.45 μM, BioRad) were incubated in pre-chilled Towbin transfer buffer (25 mM Tris, 192 mM Glycine, 0.1% (w/v)
SDS, 20% (v/v) methanol, pH 8.3) for 15 min, prior assembling the gel/membrane sandwich complex. The blotting paper (extra thick, BioRad) briefly soaked in Tobwin transfer buffer was placed on to the anode of the transfer unit, followed by the membrane, the gel and again by a pre-soaked blot paper avoiding air bubbles. For the transfer, the cathode was placed onto the complex and the transfer was done with a voltage of 10 V for 1 h. After protein transfer to the membrane, the membrane was briefly rinsed in PBS buffer (137 mM NaCl, 2.5 mM KCl, 10 mM Na₂HPO₄, 1.5 mM KH₂PO₄, pH 7.4) and then blocked with 1x PBS containing 3% (w/v) non-fat milk powder (Fluka, Sigma-Aldrich) for 1 h at room temperature. After washing the membrane three times with PBS containing 0.05% Tween 20 (Sigma-Aldrich), the membrane was incubated with PBS, 3% BSA (Fisher Scientific) and the Monoclonal Anti-polyHistidine Peroxidase Conjugate (Sigma), followed by three washes for 5 min with PBS containing 0.05% Tween 20. The membrane was developed in a solution of 2 ml of 4-Chloro-1-naphtol (Sigma-Aldrich, one tablet dissolved in 10 ml methanol), 10 ml triethanolamine buffer saline (137 mM NaCl, 27 mM KCl, 12 mM triethanolamine, pH 7.5) and 5 μl H₂O₂. The reaction was stopped by washing the membrane in water.

2.14.2 Protein detection with alkaline phosphatase conjugate

The SDS gel and the nitrocellulose membrane were prepared under the same conditions as described in Chapter 2.14.1. After the protein transfer onto the membrane, the membrane was washed in PBS buffer (137 mM NaCl, 2.5 mM KCl, 10 mM Na₂HPO₄, 1.5 mM KH₂PO₄, pH 7.4) and then incubated in PBS containing 3% BSA (Fisher Scientific) and 2% milk powder (Fluka, Sigma-Aldrich) for 60 min to block the membrane. The membrane was incubated with the primary antibody, specific for the target protein (ATR2tr and RhF reductases produced in rabbit, Covalab, 1:10000 dilution) in PBS containing 3% BSA for 60 min. The membrane was then washed twice in PBS with 0.1% Tween 20 (Sigma-Aldrich) for 5 min, followed by washes in PBS, 0.5% Tween 20 and 1 M NaCl for 5 min, rinsed briefly in PBS and again in PBS containing 3% BSA before incubating the membrane with the secondary antibody (goat Anti-rabbit IgG conjugated to alkaline phosphatase, Sigma, 1:20000 dilution) for 60 min. The membrane was then washed twice in PBS containing 0.1% Tween 20 for 5 min,
twice in PBS containing 0.5% Tween 20 and 1 M NaCl for 5 min, briefly rinsed in PBS and then incubated in 10 mM Tris pH 9.6 for 5 min. The membrane was developed with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate dipotassium (from Sigma as a tablet containing both ingredients, dissolved in 10 ml water) and the reaction was stopped by washing the membrane several times in water.

2.15 Protein characterisation

2.15.1 Characterisation of the P450 reductases

The characterisation of ATR1tr and ATR2tr was done, if not otherwise specified, with purified enzyme dialysed in buffer B: 30 mM potassium phosphate buffer (pH 7.8), 20% glycerol (Fisher Scientific), 0.1 mM EDTA (Fisher Scientific) and 20 μM FMN (Sigma-Aldrich).

2.15.1.1 Spectrophotometric characterisation

Spectrophotometric work was performed at 30 °C using a Varian Cary 50 Bio UV/Vis Spectrophotometer (Agilent Technologies). High Quality Polystyrene Disposable Cells, Semi Micro, (1.5 ml, Kartell) were used for Bradford protein quantification and UV-transparent disposable cuvette (ultra-micro, 15 mm window height, BrandTech Scientific) for activity assays and wavescan measurements. The absorption of ATR2tr (5.1 mg/ml, dialysed in 50 mM potassium phosphate buffer pH 7.5) was measured between the wavelength 300 and 600 nm with one point for each nanometre. Expressed and purified empty LIC-vector (without dialysis, protein concentration of 0.4 mg/ml) was used for negative control. FMN and FAD were dissolved in 50 mM potassium phosphate buffer pH 7.5 to a final concentration of 100 mM. The spectra of 10 μM of both substances were recorded between 300 and 600 nm.

2.15.1.2 Activity assay with cytochrome c

Activity of ATR2tr to transfer single electrons from the cofactor (NADH or NADPH) was detected by using cytochrome c as the electron donor. The protocol adapted from Guengerich et al. 2009 was modified to a 500 μl volume and the
reaction measured spectrophotometrically by following the increase in absorbance at 550 nm. Each reaction contained 50 μM horse heart cytochrome c (Sigma-Aldrich) with a certain amount of sample to a final volume of 495 μl with 300 mM potassium phosphate buffer, pH 7.7. A baseline was recorded at 550 nm for 1 min and the reaction started by adding 5 μl 10 mg/ml NAD(P)H (Melford). NADH and NADPH were tested as cofactor and ferredoxin reductase (from spinach, Sigma-Aldrich) as well as cytochrome c reductase (from porcine heart, Sigma-Aldrich) as positive control. Expressed and purified empty LIC-vector sample was used for negative control.

2.15.1.3 Temperature and pH optima of ATR1tr and ATR2tr

The temperature optimum for the two Arabidopsis reductases was determined using the activity assay with cytochrome c (Chapter 2.15.1.2) between 10 °C to 70 °C. The buffer was preheated in a water bath (Grant JB1, Scientific Laboratory Supplies) and the cuvette holder tempered.

For elucidation of the pH optimum of ATR1tr and ATR2tr, the activity for cytochrome c reduction was tested in three different buffers (Table 2.6) covering the pH range from pH 2.0-11.0. The pH optimum for both reductases was also determined in Britton Robinson buffer, which covered a pH range between pH 3.4 to 11.0.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH region</th>
<th>Content</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>potassium phosphate</td>
<td>4.0 - 10.0</td>
<td>H$_2$KPO$_4^<em>$ and HK$_2$PO$_4^</em>$</td>
<td>300 mM</td>
</tr>
<tr>
<td>citrate</td>
<td>2.0 - 7.6</td>
<td>citric acid$^\text{+}$ and tri sodium citrate$^\text{+}$</td>
<td>300 mM</td>
</tr>
<tr>
<td>Britton Robinson</td>
<td>3.4 - 11.0</td>
<td>1:1:1 mixture of boric acid$^\text{+}$, phosphoric acid$^\text{+}$ and acetic acid$^\text{+}$, pH regulated with NaOH$^\text{+}$</td>
<td>300 mM</td>
</tr>
</tbody>
</table>

*chemical from Fisher Scientific, # from Sigma-Aldrich

For the reaction, 40 μl of a 0.5 mM solution of the cytochrome c (horse heart, Sigma, dissolved in water), 1 μg/ml ATR2tr (or 100 μg/ml ATR1tr), 5 μl 10 mM NADPH (Melford) and buffer were mixed to a final volume of 500 μl (Table 4.2). The absorbance was observed at 550 nm at 30 °C. The activity (mM/min) was
calculated by using the average of three measurements divided by the extinction coefficient of cytochrome c at 550 nm ($\varepsilon = 19.6 \text{ mM}^{-1}\text{cm}^{-1}$). Each point was measured as triplicates to get a standard deviation.

2.15.1.4 Stability test of ATR2tr in different buffers

The purified ATR2tr in 500 mM imidazole was divided in two batches and one was dialysed against two 5 l volumes of buffer A (50 mM potassium phosphate buffer pH 7.5) and the other against buffer B (30 mM potassium phosphate pH 7.8, 20% glycerol, 0.1 mM EDTA, 2.0 μM FMN) at 4 °C. Buffer B containing FMN should stabilise the activity of ATR2tr due to the low affinity of ATR2 to FMN.\textsuperscript{350} Glycerol can stabilise enzymes and protect hydrophobic sites of proteins. The ATR2tr samples in buffer A or B were aliquoted and stored at -80 °C, -20 °C, 4 °C and 21 °C. The samples were assayed for activity with cytochrome c (see Chapter 2.15.1.2) after 2, 6, 16 and 41 days.

A Michaelis-Menten-diagram was generated with samples of purified ATR2tr before dialysis (containing 500 mM imidazole) and after dialysis in the two different buffers A and B using the same substrate concentrations published by Hull and Celenza 2000.\textsuperscript{351} Error bars were calculated using the standard derivation of three replicas.

2.15.1.5 Kinetic studies of the reductases

The reaction was performed at 25 °C and the absorbance change of cytochrome c was monitored at 550 nm.

The reductases ATR1tr and ATR2tr (dialysed in buffer B) were used in a final concentration of 1 mg/ml and 0.01 mg/ml, respectively. The protein solution was mixed with different cytochrome c substrate concentrations (5, 6, 10, 20, 50, 100 and 150 μM, from horse heart, Sigma) in a total volume of 495 μl 300 mM potassium phosphate buffer (pH 7.7) and the reaction was started by adding 5 μl NADPH (10 mg/ml, Melford). Measurements were taken in triplicates.

Kinetic data ($K_M$ and $V_{max}$) were calculated using Michaelis Menten kinetics (Software: GraFit) for the substrate concentration of 6 μM up to 100 μM.

The specific activity is defined by the activity of enzyme per milligram of total protein (in μmol min$^{-1}$mg$^{-1}$) and was calculated using following formula:

$\text{specific activity} = V_{max} \cdot [E]$  \hspace{1cm}  [E] = \text{enzyme concentration}$
The protein concentration of 0.01 mg and 1 mg/ml reaction for ATR1tr and ATR2tr, respectively was used for kinetic activity assays. The turnover number $k_{\text{cat}}$ is the number of moles of substrate, which are converted into product per mol of enzyme (per active site) under saturated conditions.

$$k_{\text{cat}} = \frac{V_{\text{max}}}{[E]}$$

### 2.15.2 Protein identification by MALDI-MS analysis
Protein identification was performed by the Proteomics Laboratory, Technology Facility University of York using a Bruker autoflex III MALDI-TOF/TOF and data were analysed by Mascot (Matrix Science).

### 2.15.3 P450 enzyme activity assays
A Waters (Milford, USA) HPLC system consisting of a 510 pump, a Waters Alliance 2695 separations module and a Waters 2996 Photodiode Array was used, if not otherwise specified, to detect the chemical compounds used in the different assays. The substrates and products were identified by comparing the spectra generated from the Waters photodiode array detector with the retention times and spectra of commercial compounds. Data analysis was performed using Empower-Pro Analysis Software.

#### 2.15.3.1 Activity assay for CYP71D15 with limonene
*E. coli* JM109 whole cells and the total protein fraction containing CYP71D15 PM2-2 were used for activity assay with the native substrate limonene. Resting cell assays were performed in a total volume of 1 ml containing 5 mM limonene, 0.5 mM NADPH, 20 μg/ml spinach ferredoxin and 10 U/ml spinach ferredoxin reductase (Sigma-Aldrich) and buffer (100 mM Tris-HCl (pH 7.4), 250 mM KCl, 50 mM MgCl$_2$). Samples were taken over a time course, cells spun down and then an equal volume of ethyl acetate added, vortexed, centrifuged and the organic layer analysed by GC-MS (7890A GC System, 5975C inert XL MSD with Triple-Axis Detector, 7693 autosampler, Supelco HP5 GC column, software: ChemStation) following the temperature program in Figure 2.7.
2.15.3.2 Activity assay for CYP81D8 and CYP81D11 with methyl-tolyl-sulphide

The native substrate of Arabidopsis CYP81D8 and CYP81D11 are not known, however oxidising activity was tested towards the substrate methyl-tolyl-sulphide. The reaction contained 250 µg of protein (purified with Ni-chromatography), reductase (0.1 ng/ml spinach ferredoxin and 1 U/ml spinach ferredoxin reductase (Sigma-Aldrich) or Arabidopsis ATR1tr purified see Chapter 4), 100 mM sodium phosphate buffer pH 7.0, 1 mM methyl-tolyl-sulphide and was initiated by the addition of 300 µM NADPH. The reaction (100 µl) was stopped by adding acetonitrile (30 µl) at different time points and the products were analysed by HPLC. As a positive control the heme domain of XplA was used.

HPLC was performed with 50 µl sample injected to a Techsphere column (column temperature: 45 °C, sample temperature: 4 °C, water (Elga Labwater) and methanol (Fisher Scientific), flow: 1 ml/min) following the program in Table 2.7. The retention time of the substrate methyl-tolyl-sulphide was at 14 min and of methyl-tolyl-sulphoxide at 5 min.
Table 2.7: HPLC program for the separation of methyl-tolyl-sulphide and its oxidised derivates

<table>
<thead>
<tr>
<th>time (min)</th>
<th>water (%)</th>
<th>methanol (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>5.00</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>10.00</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>10.01</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>15.00</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>15.01</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>25.00</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

2.15.3.3 Isoflavone synthase activity assay

For an *in vivo* assay in growing cell culture, the substrate naringenin (Sigma-Aldrich) was added to a final concentration of 50 μM to the growing culture at the start of induction. Flavonoids are excreted by *E. coli* and can be analysed from the culture supernatant. Therefore, 5 ml samples were collected at different time points during the expression and analysed by HPLC after extraction.

For activity tests in resting cells, cultures were harvested and resuspended in potassium phosphate buffer (50 mM, pH 7.0) with 20 mg wet cells per 1 ml buffer. After adding the substrate naringenin (Sigma, final concentration: 100 μM), the tubes were incubated at 28 °C and 160 rpm. Samples (2.5 ml) were collected at different time points for HPLC analysis after extraction.

Extraction method:

The samples were prepared according to the protocol of Leonard and Koffas:\(^\text{306}\):

To 5 ml sample the internal standard (5 μl of 50 mM scopoletin) and 5 ml ethyl acetate for extraction was added. After mixing and centrifugation (Jouan Model CR 312 Centrifuge, Swinging Bucket Rotor, SelectScience), the organic phase was transferred into a glass vial for complete evaporation. To dissolve the flavonoids, 500 μl methanol were added and the samples incubated at 60 °C for 15 min, followed by incubation at room temperature for 20 min. 50 μl were injected to a TechSphere ODS 80A 5 μ column (250 x 4.6 mm, Fischer Thermo Scientific), with an isocratic flow of 1 ml/min of 50% (v/v) methanol containing 0.1% (v/v) acetic acid and 50% (v/v) water (column temperature: 25 °C, sample temperature: 21 °C). The standards of the substrate naringenin (Sigma-Aldrich) and the product genistein (Sigma-Aldrich) eluted at 17.5 min and 12 min,
respectively. The peak area was converted into concentration using a calibration
with commercial available naringenin, genistein and scopoletin purchased from Sigma.

2.15.3.4 Cinnamate-4-hydroxylase (CYP73A5) activity assay

Cinnamate-4-hydroxylase activity was tested in resting cell assays. Therefore, the
fusion proteins were expressed in LB and M9 medium following the protocol
described in 2.10.1 (page 67) overnight at 15 °C and 200 rpm. The activity of the
73A5-fusions in the resting cell assays were optimised by variation of the
expression temperature (15 °C and 20 °C) as well as the airation (speed of 50 rpm
and 200 rpm).

Cells were harvested by centrifugation (7 min at 4000 rpm, High Speed Sorvall
RC5B+rotor, SLC-1500) and resuspended in 50 mM potassium phosphate buffer
pH 7.0 to a final cell concentration of 100 mg/ml.

The resting cell assays were performed in glass vials in a total volume of 15 ml,
the reaction started by adding the substrate cinnamic acid (Sigma-Aldrich) to a
final concentration of 200 μM. Vials were shaken (300 rpm) at 28 °C. Samples
(100 μl) were taken at different time point, and quenched with equal volume of
methanol. After centrifugation (5 min at full speed, bench top microlitre
centrifuge Sigma 1-15P), 20 μl of the supernatant were analysed by HPLC
following a modified program (Table 2.8) according to Chen and Morgan353 using
a Techsphere column (column temperature: 30 °C, sample temperature: 4 °C,
buffer: water (Elga Labwater) and methanol (Fisher Scientific) containing 0.1%
(v/v) acetic acid (Fisher Scientific) buffered to pH 7.25 with triethylamine
(Sigma-Aldrich). Cinnamic acid (absorption maximum: 278 nm) eluted at 9.4 min
and coumaric acid (absorption maximum: 310 nm) at 7.7 min under these
conditions.
Chapter 2 – Materials and Methods

Table 2.8: HPLC program for the separation cinnamic acid and coumaric acid at 1 ml/min

<table>
<thead>
<tr>
<th>time (min)</th>
<th>water (%)</th>
<th>methanol containing 0.1% acetic acid, pH 7.25 with triethylamine (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>3.00</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>4.00</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>9.00</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>9.01</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>14.00</td>
<td>90</td>
<td>10</td>
</tr>
</tbody>
</table>

Ni-affinity chromatographic treated 73A5tr-ATR2tr (expressed in M9 at 15 °C and 200 rpm and purified in a continued process, see Chapter 2.10, page 67), protein bound to the resin material (0.5 ml resin material after two washing steps) as well as the insoluble protein (700 mg wet cells after sonication) were used for crude extract assays. The reactions were performed in glass vials in a total volume of 4 ml containing 4 U alcohol dehydrogenase (ADH, from *T. brockii*, Sigma-Aldrich), 2.5% (v/v) isopropanol (Fisher Scientific, substrate for ADH), 500 μM substrate cinnamic acid and 300 μM NADPH (Melford) shaking (300 rpm) at 30 °C. Samples were taken and prepared as described for the resting cell assay (see above).

A standard curve was measured for cinnamic acid and coumaric acid in the range of 10 μM to 500 μM with four replicas for each concentration.

2.15.3.5 Activity assay for CYP82E4

The *N*-demethylase (CYP82E4) from *Nicotiana tabacum* converts nicotine to nornicotine, both substances are possible to detect by HPLC analysis according to Saunders and Blume 1981. Samples (sample temperature: 4 °C) were separated on a SunFire C18 column (3.5 μm, 4.6 x 150 mm, Waters, column temperature: 20 °C), 1 ml/min flow of isocratic mobile phase of 70% water containing 0.2% phosphoric acid (pH 7.25 with triethylamine) and 30% methanol containing 0.2% phosphoric acid (pH 7.25 with triethylamine) with a run time of 28 min. The retention time of substrate (±) nicotine (Sigma) was at 18.9 min and of the product (±) nornicotine (Sigma) at 4.2 min (Figure 2.8). Both have an absorption maximum of 260 nm.
CHAPTER 2 – MATERIALS AND METHODS

The activity for the fusions 82E4tr-lamATR2tr and 82E4tr-licATR2tr (sequences see appendix 2) was tested in a resting cell assay. The assay was performed in glass vials containing 25 mg cells per ml 50 mM potassium phosphate buffer (pH 7.0) and 100 μM substrate nicotine in a final reaction volume of 20 ml, shaking at 28 °C and 300 rpm. Samples of 1.2 ml were taken at different time points, centrifuged (5 min at full speed, bench top microlitre centrifuge Sigma 1-15P) and 1 ml of the supernatant was evaporated completely (Savant SpeedVac DNA 110 Concentrator). The pellet was resuspended in 100 μl water : methanol (70% : 30%) and 50 μl were analysed by HPLC.

Additionally, HPLC separation of 1 mM nicotine and 1 mM nornicotine was tested with a ChiralPak IA column (Daicel Chemical Industries). Compounds were separated in a mobile phase of 35% (v/v) ethanol and 65% (v/v) (heptane : isopropanol : TFA =90:10:0.1%) with a isocratic flow of 0.5 ml/min and a column temperature of 20 °C. The substances were detected at 260 nm.

2.15.3.6 Activity assay for CYP81D8 with TNT and aminodinitrotoluenes

The resting cell assay against TNT was performed with CYP81D8tr, 81D8tr-lamATR2tr, 81D8tr-licATR2tr (all cloned into the LIC-vector) in 50 mM potassium phosphate buffer (pH 7.0) after an expression in E. coli Rosetta 2 (DE3) in M9 medium (2.10.1, page 67). For negative controls lamATR2tr,
licATR2tr (both in LIC-vector) and Rosetta 2 (DE3) cells without a plasmid were used.

The total reaction volume of 5 ml contained 50 mg wet cells per ml buffer (50 mM potassium phosphate buffer, pH 7.0). TNT (stock: 100 mM TNT in DMSO) was added to a final concentration of 100 μM and samples incubated at 28 °C shaking at 200 rpm. Samples were taken over a time course, quenched by adding 10% of 1.5 M trichloroacetic acid (TCA), snap frozen in liquid nitrogen and stored at -80 °C, due to the instability of hydroxylaminodinitrotoluene in presence of oxygen. Defrosted samples were centrifuged (10 min at full speed, bench top microlitre centrifuge Sigma 1-15P) and 50 μl applied for HPLC analysis. For the separation of the TNT derivates a TechSphere ODS 80A 5 μm column (250 x 4.6 mm, Fischer Thermo Scientific) was used. 50 μl of each sample were run in the solvent conditions of 40% (v/v) water and 60% (v/v) methanol with a flow of 1 ml/min for 10 min (column temperature: 35 °C, sample temperature: 4 °C). TNT eluted after 5.6 min and the derivates 4-hydroxylamino-2,6-dinitrotoluene (4-HADNT) at 4.7 min, 2-amino-4,6-dinitrotoluene (2-ADNT) at 6.1 min, 4-amino-2,6-dinitrotoluene (4-ADNT) at 5.6 min, 2,4-dinitrololuene at 6.7 min and 2,6-dinitrololuene at 7.2 min.

2.15.3.7 Activity assay for CYP81D8 with 7-ethoxycoumarin

Recombinant expressed CYP81D8 (in E. coli Rosetta 2 (DE3) and M9 medium, Chapter 2.10.1, page 67) 81D8tr-lamATR2tr and 81D8tr-licATR2tr (negative controls: control-lamATR2tr, control-licATR2tr and Rosetta 2 (DE3) cells without a plasmid) were tested in a resting cells assay with 7-ethoxycoumarin, which is a common P450 substrate.

Therefore, 50 mg cells per ml buffer (50 mM potassium phosphate buffer, pH 7.0) and 200 μM 7-ethoxycoumarin (Sigma-Aldrich) in a final volume of 5 ml were incubated in glass vials at 28 °C shaking at 200 rpm. Samples (300 μl) were taken over a time course, centrifuged (10 min at full speed, bench top microlitre centrifuge Sigma 1-15P) to remove the cells. Fifty μl of the supernatant were analysed by HPLC (Waters 717 Plus Autosampler, 2487 dual λ absorbance detector, Waters SunFire C18 column, 3.5 μm, 4.6 x 150 mm) at a wavelength of 325 nm under isocratic 50% (v/v) water and 50% (v/v) methanol with 0.1% (v/v) acetic acid, flow: 0.7 ml/min (run time: 35 min, column temperature: 21 °C,
sample temperature: 4 °C). The retention times of 7-ethoxycoumarin and 3-hydroxycoumarin (Sigma-Aldrich), a possible product, were 18.2 and 10.2 min, respectively.
3 Chapter: Expression of Cytochromes P450

3.1 Introduction

Hosts such as yeast (eukaryotic host) or bacteria (prokaryotic host) have been traditionally used to express recombinant P450s for characterisation studies and for industrial applications. Both have advantages and disadvantages: The yeast expression is efficient and inexpensive and can be up-scaled for industrial applications; moreover, it allows expression of native membrane associated P450s in microsomes. The yeast *Saccharomyces cerevisiae* possesses three endogeneous P450s and one reductase representing a potential disadvantage of this system as endogeneous P450s could interfere with the P450 of interest.

The bacterium *E. coli* lacks P450s and can be more easily, faster and more economically cultivated in the lab. Expression problems can occur through different codon bias and protein misfolding due to the absence of complex membrane systems like the ER and post-translational machinery.

*S. cerevisiae* contains an endogenous P450 reductase which is able to efficiently transfer electrons to foreign P450s.\(^{323}\) *S. cerevisiae* was the first host used to successfully express mammalian\(^{324,325}\) and plant\(^{326}\) P450s with activity of membrane associated proteins being detected in the microsomes. A *S. cerevisiae* strain has been specially modified by introducing the Arabidopsis P450 reductase ATR1 and ATR2 genes and silencing the endogeneous yeast P450 reductase. These strains were named WAT11 and WAT21 (Table 3.1).\(^{166,338}\) More than 20 Arabidopsis P450s have been expressed using these engineered yeast strains (Chapter 1.5.2).

<table>
<thead>
<tr>
<th><em>S. cerevisiae</em></th>
<th>genotype</th>
<th>obtained from</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAT11</td>
<td>MAT(a); ade2–1; his3–11,-15; leu2–3,-112; ura3–1; can(^R); cyr(^+) (a derivative of the W303-B strain)</td>
<td>Prof. D. Werck Reichhart (CNRS-Institute de Biologie moléculaire des plantes, Strasbourg)</td>
</tr>
<tr>
<td>WAT12</td>
<td>MAT(a); ade2–1; his3–11,-15; leu2–3,-112; ura3–1; can(^R); cyr(^+) (a derivative of the W303-B strain)</td>
<td></td>
</tr>
</tbody>
</table>
E. coli was also used as host for an expression of membrane-associated Arabidopsis P450s such as CYP74A1, 79A2, 79B2, 79B3, 79F1, 90B1, 98A3 and 701A3, however so far, Arabidopsis P450s have been not expressed solubly in E. coli.

Activity of recombinant P450s has mostly been detected using radiolabelled substrates. Other methods used for activity detection include HPLC, LC-MS, GC-MS and Solid State Nuclear Magnetic Resonance (SSNMR).

The two Arabidopsis P450s CYP81D8 and CYP81D11 were selected for this project as these enzymes were upregulated in response to TNT (2,4,6-trinitrotoluene, Figure 3.1) and might play a role in general detoxification via a broad substrate specificity (Table 3.2).

![Figure 3.1: Chemical structure of 2,4,6-trinitrotoluene (TNT)](image)

<table>
<thead>
<tr>
<th>Fold increase</th>
<th>Accession number</th>
<th>Gene family name</th>
</tr>
</thead>
<tbody>
<tr>
<td>30.7</td>
<td>At4g37370</td>
<td>CYP81D8</td>
</tr>
<tr>
<td>25.1</td>
<td>At3g28740</td>
<td>CYP81D11</td>
</tr>
<tr>
<td>14.5</td>
<td>At5g57220</td>
<td>CYP81F2</td>
</tr>
<tr>
<td>13.1</td>
<td>At2g30750</td>
<td>CYP71A12</td>
</tr>
<tr>
<td>8.7</td>
<td>At3g26830</td>
<td>CYP71B15</td>
</tr>
<tr>
<td>5.8</td>
<td>At5g67310</td>
<td>CYP81G1</td>
</tr>
<tr>
<td>3.4</td>
<td>At5g45340</td>
<td>CYP707A3</td>
</tr>
<tr>
<td>2.3</td>
<td>At3g26210</td>
<td>CYP71B23</td>
</tr>
</tbody>
</table>

The explosive TNT is highly toxic, recalcitrant to degradation in the environment and a major world-wide pollutant of our environment. This modern, synthetic compound is used for military purposes and also for the demolition industry. The TNT molecule is not a natural compound and has existed in the environment for a
relatively short time. These factors are likely to have contributed to the fact that there are not many organisms able to transform TNT to a less toxic compound. It is known, using microarray analyses, that expression of genes encoding enzymes from Arabidopsis, such as uridine diphosphate glycosyltransferases (UGTs), oxyphytodienoate reductases (OPRs), glutathione transferases (GSTs) and P450s, are upregulated in response to TNT stress. UTGs, OPRs and GSTs have been shown to have activity towards TNT and its metabolites. Microarray data available on the Internet (Genvestigator, www.genevestigator.com) were used to find out details on the expression profiles of the two P450s chosen from the microarray experiment. CYP81D8 is expressed predominantly in the seedling and the young rosette and again in the mature silique (Figure 3.2). CYP81D11 is mainly produced during and after germination and in the first stage of rosette growth. Transcripts of CYP81D11 are less abundant in the later stages of plant development.

![Graph showing expression levels of CYP81D8 and CYP81D11](https://www.genevestigator.com/gv/, 15/12/2011)

**Figure 3.2:** Occurrence of CYP81D8 and CYP81D11 in the development of Arabidopsis

(https://www.genevestigator.com/gv/, 15/12/2011)
Another P450, the (−)-4S-limonene-3-hydroxylase (CYP71D15) from peppermint (Mentha x piperta), a hybrid of spearmint (Mentha spicata) and watermint (Mentha aquatica), was studied in this chapter, because the limonene activity have been measured in vitro by GC-MS after solubilisation of the membrane associated protein CYP71D15.  

CYP71D15 hydroxylates (−)-4S-limonene in the C3 position regio- and stereospecifically to (−)-trans-isopiperitenol, which is a precursor of (−)-menthol (Figure 3.3).  

(−)-Menthol is a well studied monoterpene that occurs naturally as essential oil in peppermint. It is used in the food industry, for oral health care, cosmetics and tobacco due to its pleasant aroma and flavour and the cooling-anaesthetic effect.  

Prof. Rodney Croteau’s research group studied the limonene-3-hydroxylase after recombinant expression in E. coli JM109 and in the yeast strains WAT11 and WAT21. Hydroxylation rates of more than 2000 nmol·h⁻¹·mg⁻¹ by the membrane associated CYP71D15 were achieved after expression in E. coli, whereas only 35 nmol·h⁻¹·mg⁻¹ were obtained with yeast microsomes, independent of the strain (WAT11 or WAT21).  

The hydrophobic membrane anchor of P450s and its reductase can be removed to increase the solubility with no activity loss as suggested in the literature. Kempf et al. reported that the human CYP2D6 was expressed solubly in the cytosol after removing the hydrophobic membrane anchor.  

Figure 3.3: Regiospecific hydroxylation of (−)-4S-limonene to (−)-trans-isopiperitenol within the biosynthesis of (−)-menthol
3.2 Objectives

Different strains of *E. coli* as well as *S. cerevisiae* were tested as hosts for the expression of the plant P450s: CYP81D8 and CYP81D11, both from Arabidopsis and the peppermint CYP71D15. Therefore, the native gene sequences were expressed in *S. cerevisiae* and then tested for activity. Truncated versions (without the hydrophobic membrane anchor to increase the solubility) of the P450s were used for expression in *E. coli*.

*E. coli* expression was carried out for the three P450s and the expression yield optimised by testing a range of strains and experimental conditions.

The yeast work was done in York and in Prof. Danièle Werck-Reichhart’s laboratory (Institute de Biologie Moléculaire des plants du CNRS, Département Réponses au Stress, Strasbourg, France). Two modified yeast strains, WAT11 and WAT21338 for coexpression of the Arabidopsis reductases ATR1 and ATR2, respectively, were provided by Prof. Danièle Werck-Reichhart for the expression of full length plant P450s.

3.3 Materials and Methods

3.3.1 Analysis of CYP73D15 PM2-2 construct

Prof. Rodney Croteau (Institute of Biological Chemistry, Washington State University, Pullman, US) provided the CYP71D15 in the vector pCWori+ (construct PM2-2367). The amino acid sequence of the N-terminus in the construct PM2-2 was altered incorporating seven residues of a bovine 17α-hydroxylase (Figure 3.4).321,367 A vector map is shown in Figure 3.5.

<table>
<thead>
<tr>
<th>native CYP71D15</th>
<th>MELLOQWSALIILV</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM2-2</td>
<td>MALLLA VFWSALIILV</td>
</tr>
<tr>
<td>native CYP71D15</td>
<td>ATGGAGCTCCTCAGCTTTTGTCGGCGCTTTATAATCCTCGTAG</td>
</tr>
<tr>
<td>PM2-2</td>
<td>ATGGCTCGTTATAGCAGTTTTTGGTCGGCGCTTTATAATCCTCGTAG</td>
</tr>
</tbody>
</table>

*Figure 3.4: N-terminal modification of CYP71D15 creating construct PM2-2*367
Figure 3.5: Vector map for pCWori+
LacZ = β-galactosidase gene, LacI = repressor gene for IPTG induction, tac = tac promoter, CYP71D15 PM2-2 = gene of interest, M13 ori = origin of replication of filamentous phage M13, ampR = ampicillin resistance gene, NdeI and HindIII = restriction sites

The CYP71D15 PM2-2 insert was analysed by a restriction digest using NdeI and HindIII and sequenced using primers listed in Table 3.3 (Chapter 2.8.9).

Table 3.3: Primer used for sequence analysis of CYP71D15 PM2-2

<table>
<thead>
<tr>
<th>primer</th>
<th>sequence (5’→ 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tac promoter F</td>
<td>CCATGGTGACAATTAATCATCGGCTC</td>
</tr>
<tr>
<td>PM2-2int1263 F</td>
<td>GGGGAATGATTTTCGAGTTTCGTCCCCGTTCGG</td>
</tr>
<tr>
<td>Pm2-2int242 R</td>
<td>GGAGAACACCTCACCCAGCTGCAG</td>
</tr>
<tr>
<td>CYP71D15 PM2-2 R</td>
<td>GAGGAGAAGGCAGTGCATGATGAAAGATCGAGGGTGTGGGAAC</td>
</tr>
</tbody>
</table>
3.3.2 Materials and Methods for P450 expression in *Escherichia coli*

3.3.2.1 Isolation and cloning of the plant P450s

RNA was isolated from 14 days old Arabidopsis plants grown in liquid medium treated with TNT (Chapter 2.7.2) and transcribed into cDNA (Chapter 2.7.3). Usually, P450s have a hydrophobic N-terminal region, allowing association with the membrane. The P450 amino acid sequences were analysed using the software TMHMM and SignalP3.0 (Chapter 2.8.1). The hydrophobic N-termini were removed by designing special LIC primers (Chapter 1.1) for the PCR amplification of Arabidopsis CYP81D8 and CYP81D11 (Primer see Table 3.4).

Table 3.4: Primer used for PCR amplification of the truncated versions of CYP81D8 and CYP81D11

<table>
<thead>
<tr>
<th>primer</th>
<th>sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP81D8nat F</td>
<td>CCAGGGGACCAGCAATGGGAAAACCAAAACCCTAAATTTTCCTCATTTCTCC</td>
</tr>
<tr>
<td>CYP81D8tr F</td>
<td>CCAGGGGACCAGCAATGGGAAAACCAAAACCCTAAATTTTCCTCATTTCTCC</td>
</tr>
<tr>
<td>CYP81D8 R</td>
<td>GAGGAGAAGGGCGCGTCAAACGGGACTCGTTGAAAGATTATTTAAACACAGAGG</td>
</tr>
<tr>
<td>CYP81D11nat F</td>
<td>CCAGGGACCAGCAATGCTCATTACAAACAAAAAGCAATAATGGAAACTATATACC</td>
</tr>
<tr>
<td>CYP81D11tr F</td>
<td>CCAGGGGACCAGCAATGCTCATTACAAACAAAAAGCAATAATGGAAACTATATACC</td>
</tr>
<tr>
<td>CYP81D11 R</td>
<td>GAGGAGAAGGGCGCGTCAAACGGGACTCGTTGAAAGATTATTTAAACACAGAGG</td>
</tr>
<tr>
<td>CYP71D15 PM2-2 F</td>
<td>CCAGGGACCAGCAATGCTCATTACAAACAAAAAGCAATAATGGAAACTATATACC</td>
</tr>
<tr>
<td>CYP71D15 PM2-2 R</td>
<td>GAGGAGAAGGGCGCGTCAAACGGGACTCGTTGAAAGATTATTTAAACACAGAGG</td>
</tr>
</tbody>
</table>

The peppermint CYP71D15 PM2-2 and the two truncated versions of Arabidopsis P450s (81D8tr and 81D11tr) were amplified using Phusion polymerase (Chapter 2.8.2) and primer listed in Table 3.4 before cloning into the LIC-vector (Chapter 1.1).

3.3.3 Expression of PM2-2 in *Escherichia coli*

For microsome preparation\(^{367,372}\) *E. coli* JM109 cells transformed with CYP71D15 PM2-2 were grown with shaking (200 rpm) at 37 °C in TB (Chapter 2.2.1) containing 100 μg/ml ampicillin. Expression was induced at OD\(_{600nm}\) = 0.85 by adding 1 mM IPTG, 1 mM thiamine and 75 μg/ml ALA (*E. coli* JM109 without the vector was used as negative control) and samples incubated with shaking (200 rpm) at 28 °C following the published protocol.\(^{367}\) After 40 h the cells (800 ml) were harvested by centrifugation (5000 rpm, rotor:
SLC1500, 10 min at 4 °C), resuspended in 35 ml 100 mM Tris buffer (pH 7.5) containing 20% glycerol, 0.5 M EDTA, 1 mM DTT and 0.2 mg/ml lysozyme. Cells were incubated with stirring for 10 min at 4 °C. Cell were centrifuged again (7000 rpm, rotor: SLC1500, 10 min at 4 °C) and resuspended in 30 ml 100 mM Tris buffer (pH 7.5) containing 20% glycerol, 0.5 M EDTA, 1 mM DTT and PMSF. Cells were sonicated (Chapter 2.10.3) and centrifuged at low speed (SS34-rotor, 3250 rpm). The supernatant was then spun in an ultracentrifuge (40 krpm for 1 h at 4 °C, rotor type 45Ti) to separate the soluble proteins from the membrane fraction. Membrane proteins were resuspended in 100 mM sodium phosphate buffer (pH 7.5) containing 30% glycerol using a glass homogeniser and proteins were detected by SDS-PAGE and western blot analysis.

For soluble, non membrane associated expression of the truncated Arabidopsis P450s 81D8tr and 81D11tr as well as the peppermint CYP71D15 PM2-2 several different strains (DH5α, BL21(DE3), Rosetta 2, Rosetta gami 2) and media (LB, TB, M9 and auto-induction) was trialled. Incubations were all performed overnight with shaking (200 rpm) at 20 °C following the protocol Chapter 2.10.1. The purification was done in a batch process (Chapter 2.10.3-2.10.4).

### 3.3.4 Materials and Methods for P450 expression in yeast

#### 3.3.4.1 Media

For yeast cultivation YPAG (yeast, bactopetone, adenine and glucose) medium was used containing 10 g/l yeast extract (Formedium), 10 g/l bacto peptone (Difco), 20 g/l glucose (Fisher Scientific) and 200 mg/l sterile filtrated adenine was added after sterilisation. SGI medium (glucose based yeast minimal medium) was used for the selection of transformants on agar containing 1 g/l bacto casamino acids, 7 g/l yeast nitrogen base without amino acids (Difco), 20 g/l glucose, 20 g/l agar (Formedium) and 20 mg/l sterile tryptophane (Sigma-Aldrich) was added after sterilisation. The cultivation of the transformants in liquid culture was done in YPGE (yeast, bactopetone, glucose and ethanol): 10 g/l yeast extract, 10 g/l bactopeptone and 5 g/l glucose. Ethanol (50.6 ml/l) was added after autoclaving.
3.3.4.2 Transformation and expression in yeast

CYP71D15 from peppermint, CYP81D8 and CYP81D11 from Arabidopsis were cloned into the pYeDP60 vector. The vector pYeDP60 (vector map see Figure 2.3) is a shuttle vector for cloning in *E. coli* and expression in yeast. The transformation was carried out following the protocol adapted from Schiestl & Gietz, 1989. Briefly, 50 ml YPGA liquid preculture was inoculated with an isolated yeast colony and incubated with shaking (140 rpm) at 28 °C over night. Then the main YPGA culture was inoculated to an OD_{700nm} 0.2 and incubated for five hours with shaking (140 rpm) at 28 °C. Cells were then centrifuged at low speed (500x g) for 10 min, the pellet resuspended in sterile water (1/10 of the original culture volume), transferred in 2 ml Eppendorf tubes (1 ml/tube) and centrifuged at low speed (500x g) for 10 min. The pellet was resuspended in 1.5 ml 0.1 M LiAc/TE (0.1 M LiAc, 10 mM Tris-HCl, 1 mM EDTA), centrifuged at low speed (500x g for 10 min) and the supernatant discarded leaving approximately 50 μl yeast cell suspension. To this yeast cell suspension, 1-5 μg plasmid-DNA, 10 μl carrier DNA solution (100 μl 10 mg/ml salmon sperm DNA in sterile water denatured for 20 min at 100 °C and then chilled on ice) and 40% PEG in 0.1 M LiAc/TE were added and cells were incubated gently shaking for one hour at 30 °C. After heat shock at 42 °C for 15 min, the suspension was centrifuged for 10 sec at low speed, the pellet washed with 1 ml sterile water, centrifuged again and resuspended in 200 μl SGI and plated on SGI agar. Agar plates were incubated at 30 °C for 3-4 days until colonies appeared.

For the expression, 10 ml SGI liquid preculture were inoculated with white transformants (colonies without a plasmid have a lack in adenine and turn red) and incubated at 28 °C with shaking (140 rpm) over night. The main culture was done in 6 x 200 ml YPGE, which was inoculated with 1.2-2 ml preculture and incubated with shaking (140 rpm) at 28 °C for 24 hours. After this time the glucose is used up by the yeast and it is using now the ethanol for its metabolism until 20 ml galactose (200g/l) were added for induction to each flask and incubated shaking (140 rpm) for 16 hours at 25 °C. Then the yeast was harvested and the microsoms isolated (see 3.3.4.3).
3.3.4.3 Yeast microsome preparation

After the expression, the yeast cells were centrifuged for 15 min at 7500 g at 4 °C and washed with TEK buffer (1 ml TEK/0.5 g yeast, TEK: 50 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM KCl). The pellet was washed twice with 1 ml TES/0.5 g yeast (TES: 50 mM Tris-HCl pH 7.5, 1 mM EDTA, 600 mM sorbitol, fresh added 10 g/l BSA and 120 µl β-mercaptoethanol), glass beads (glass beads, 0.40-0.60 mm, Sartorius) were added to a fifth of the volume and shaken for 5 min at 4 °C (1 min shaking, 1 min chilled on ice). Glass beads were allowed to settle and the supernatant was transferred into fresh tubes. Glass beads were then washed twice with TES and the washings combined with the supernatant. The solution containing the shared cells was centrifuged (15 min at 7500x g at 4 °C) for removing cell membranes, leaving membrane fractions and soluble proteins in the supernatant. The supernatant was filtered with micracloth (Calbiochem, Merck) to remove all glass beads and then ultracentrifuged at 100000x g at 4 °C (60 Ti-32000 rpm, 45 Ti-30000 rpm) for 45 min to pellet the microsomes. The pellet was transferred into a potter homogeniser, 500 µl TEG (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 30% glycerol) added and homogenised on ice avoiding the build up of air bubbles.

To test the presence of active P450, sodium dithionite (Fluka, Sigma-Aldrich) and CO was added by bubbling the cuvette for 30 sec to create the CO bound form of the P450. Absorption was measured between 400 and 500 nm. Purified XplA heme$^{352}$ was used as positive control.

3.4 Results

3.4.1 P450 amino acid sequence analysis

Plant P450s have a membrane anchor of around 25 amino acids at the N-terminal sequence, which is hydrophobic and known to cause the insolubility of these proteins. The catalytic centre is situated in the cytoplasm, so that a deletion of the N-terminus should not reduce the enzyme activity.

The amino acid sequence was analysed with SignalP and TMHMM software (Chapter 2.8.1, result in Figure 3.6) and primers designed excluding the first 22 amino acids from the N-terminus in order to clone 81D8tr without the membrane
anchor. To ensure expression an ATG start codon was added to the 5’end of the forward primer.

**Figure 3.6: Analysis of the CYP81D8 amino acid sequence**

A SignalP blot (cleavage prob = probability of cleavage site, n-region prob = probability of N-terminus of the signal peptide, h-region prob = probability of hydrophobic region of signal peptide, c-region prob = probability of C-terminus of signal peptide) and B TMHMM
A similar approach was taken to the cloning of CYP81D11. Here the first 32 amino acids (Figure 3.7) were removed to expressed 81D11tr without a membrane anchor. Again an ATG start codon was added to the 5’end of the forward primer.

**Figure 3.7: Analysis of the CYP81D11 amino acid sequence**

A SignalP blot (cleavage prob = probability of cleavage site, n-region prob = probability of N-terminus of the signal peptide, h-region prob = probability of hydrophobic region of signal peptide, c-region prob = probability of C-terminus of signal peptide) and B TMHMM
The analysis result of the N-terminal amino acid sequence of CYP71D15 PM2-2 (Figure 3.8) was similar to the two Arabidopsis P450s described above. The full length of the CYP71D15 PM2-2 was used due to the fact that the N-terminus of the construct PM2-2 was already optimised for *E. coli* expression by Haudenschild *et al.* 2000\(^{367}\).

**Figure 3.8: Analysis of the CYP71D15 amino acid sequence**

A SignalP blot (cleavage prob = probability of cleavage site, n-region prob = probability of N-terminus of the signal peptide, h-region prob = probability of hydrophobic region of signal peptide, c-region prob = probability of C-terminus of signal peptide) and B TMHMM
3.4.2 Analysis of CYP71D15 PM2-2

The pCWori+ vector (5.5 kb) containing the insert CYP71D15 PM2-2 (1.5 kb) was digested and fragments separated using an agarose gel (Figure 3.9).

Different internal and external primers were used for sequencing the PM2-2 gene and flanking regions. Sequencing confirmed the CYP71D15 PM2-2 gene and C-terminal His-Tag comprising four histidine residues. The complete nucleotide sequence is found in Appendix B.
3.4.3 P450 cloning into the LIC-vector

All three P450s were amplified by PCR using Phusion polymerase (Figure 3.10) and then cloned into the LIC-vector.

![Image of gel electrophoresis showing four lanes labeled 8nat, 8tr, 11nat, and 11tr, with bands at different lengths.](image.png)

Figure 3.10: Analysis of PCR products for native CYP81D8 (8nat), N-truncated CYP81D8 (8tr), native CYP81D11 (11nat) and N-truncated CYP81D11 (11tr)
3.4.4 **P450 expression in *Escherichia coli***

The peppermint CYP71D15 PM2-2 was expressed in *E. coli* JM109 following the published protocol from Prof. Rodney Croteau’s research group (Chapter 3.3.3). SDS-PAGE analysis was performed to verify the expression of CYP71D15 PM2-2. The expected size of CYP71D15 PM2-2 was 57 kDa, however no signal for CYP71D15 PM2-2 was detected (Figure 3.11).

![Figure 3.11: Analysis of crude extract of JM109 cells after expression of CYP71D15 PM2-2 (57 kDa)](image)

JM109 cells without plasmid was used for negative control (1 = cells before induction, 2 = cells after induction, 3 = disrupted cells after sonication, 4 = supernatant after sonication, 5 = soluble protein after ultracentrifugation, 6 = insoluble protein fraction containing membrane proteins after ultracentrifugation)

The different fractions were also analysed using a spectrophotometer, but the typical shift of the absorption maximum from 420 nm to 450 nm when bound carbon monoxide was absent.

All three P450s were expressed in different *E. coli* strains (DH5α, BL21(DE3), Rosetta 2, Rosetta gami 2) and various media were tested. The results are summarised in Table 3.5.
Table 3.5: Scheme of the test expression of the Arabidopsis P450s in the Lic-vector using different *E. coli* strains and media

<table>
<thead>
<tr>
<th></th>
<th>LB-medium</th>
<th>TB-medium</th>
<th>M9-medium</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CYP71D15</strong></td>
<td>no expression</td>
<td>no expression</td>
<td>not tested</td>
<td>BL21(DE3)</td>
</tr>
<tr>
<td><strong>PM2</strong></td>
<td>no expression</td>
<td>no expression</td>
<td>not tested</td>
<td>Rosetta 2</td>
</tr>
<tr>
<td><strong>81D8tr</strong></td>
<td>no expression</td>
<td>not tested</td>
<td>no expression</td>
<td>Rosetta gami 2</td>
</tr>
<tr>
<td><strong>expression</strong></td>
<td>expression</td>
<td>not tested</td>
<td>no expression</td>
<td>BL21(DE3)</td>
</tr>
<tr>
<td><strong>no expression</strong></td>
<td>no expression</td>
<td>not tested</td>
<td>not tested</td>
<td>Rosetta gami 2</td>
</tr>
<tr>
<td><strong>81D11tr</strong></td>
<td>no expression</td>
<td>not tested</td>
<td>not tested</td>
<td>DH5α</td>
</tr>
<tr>
<td><strong>expression</strong></td>
<td>expression</td>
<td>not tested</td>
<td>no expression</td>
<td>Rosetta 2</td>
</tr>
<tr>
<td><strong>no expression</strong></td>
<td>no expression</td>
<td>not tested</td>
<td>not tested</td>
<td>Rosetta gami 2</td>
</tr>
<tr>
<td><strong>no expression</strong></td>
<td>no expression</td>
<td>not tested</td>
<td>not tested</td>
<td>DH5α</td>
</tr>
</tbody>
</table>

Expression was detected for 81D8tr (54 kDa) and 81D11tr (53 kDa) when expressed in *E. coli* Rosetta 2 and LB medium (Figure 3.12 and Figure 3.13). Uninduced culture and induced empty vector culture were used as negative control.

![Figure 3.12: SDS PAGE analysis of crude extract from E. coli Rosetta 2 cells expressing 81D8tr (54 kDa) in LB medium with and without induction](image)

An increase in signal could be seen from 3 h on.
Figure 3.13: SDS PAGE analysis of crude extract from E. coli Rosetta 2 cells expressing 81D11tr (53 kDa) in LB medium with and without induction
An increase in signal could be seen from 3 h on.

Additionally auto-induction medium was tested and P450s were purified in a single step using Ni-affinity chromatography. Fractions were analysed using SDS PAGE gel electrophoresis. No signal was detected for 81D8tr and 81D11tr in whole cells, disrupted cells, soluble or insoluble protein fractions, possibly due to low expression levels. However, a strong signal was found in the eluted protein fractions for both P450s (Figure 3.14).

Figure 3.14: SDS PAGE analysis of protein fractions containing 81D8tr and 81D11tr after purification on Ni-resin material
A 81D8tr (54 kDa) and B 81D11tr (53 kDa) after expression in E. coli Rosetta gami 2 in auto-induction medium (1 = whole cells before sonication, 2 = disrupted cells, 3 = soluble protein fraction, 4 = insoluble protein fraction, 5 = eluted protein fraction)
A similar result was seen after purification of 81D8tr, 81D11tr (Figure 3.15). No signal was detected for 71D15 PM2-2 (57 kDa, Figure 3.16A) when expressed in *E. coli* Rosetta gami 2 and LB medium. Cells harbouring the empty were used as negative control and purified as the P450s (Figure 3.16B).

Figure 3.15: SDS PAGE analysis of protein fractions containing 81D8tr and 81D11tr after purification on Ni-resin material

A 81D8tr (54 kDa) and B 81D11tr (53 kDa) after expression in *E. coli* Rosetta gami 2 in LB medium (1 = whole cells before sonication, 2 = disrupted cells, 3 = soluble protein fraction, 4 = insoluble protein fraction, 5 = eluted protein fraction)

Figure 3.16: SDS PAGE analysis of protein fractions containing CYP71D15 PM2-2 and empty vector control after purification on Ni-resin material

A CYP71D15 PM2-2 (57 kDa) and B empty vector control after expression in *E. coli* Rosetta gami 2 in LB medium (1 = uninduced cells, 2 = disrupted cells, 3 = soluble protein fraction, 4 = insoluble protein fraction, 5 = dialysed protein fraction)
Additionally a western blot analysis was performed to check if there were proteins possessing a His-Tag, but no signal was detected for any of the three P450s (Figure 3.17).

![Western Blot Analysis](image)

**Figure 3.17: Analysis of plant P450s after purification on Ni-resin material**

A. SDS-PAGE and B. western blot of purified P450s (1 = 81D8tr, 2 = 81D11tr, 3 = potassium phosphate buffer negative control, 4 = CYP71D15 PM2-2)

### 3.4.5 Activity assay for 81D8tr, 81D11tr and CYP71D15 PM2-2

P450 activity was tested in resting cell assays, disrupted (sonicated) cells and Ni-chromatography purified protein although P450 expression could be not detected by SDS PAGE and western blot analysis, possibly due to too low expression.

No native substrate is known for the two Arabidopsis CYP81D8 and CYP81D11. Thus to test the ability of these P450s to perform an oxidative reaction methyl-tolyl-sulphide was tested as substrate. The bacterial P450 XplA heme domain (with ferredoxin and ferredoxin reductase from spinach as reductase part) can oxidise methyl-tolyl-sulphide to methyl-tolyl-sulphoxide and was used as positive control.

For the bacterial P450 XplA the substrate peak decreased and a corresponding peak for methyl-tolyl-sulphoxide appeared (Figure 3.16). For all other tested samples, the concentration of the volatile methyl-tolyl-sulphide substrate decreased over time, but no product peak was detected. Therefore, it is possible that a reaction has occurred, however, not producing the predicted reaction product methyl-tolyl-sulphoxide. Additionally, whole cells and total protein
fraction were tested for activity against methyl-tolyl-sulphide, but no activity could be detected (results not shown).

Figure 3.18: HPLC chromatogram of the methyl-tolyl-sulphide assay
(black line = positive control at time zero)

CYP71D15 PM2-2 was tested with its native substrate limonene using whole cells and total protein. The substrate limonene (10 μg/ml) had a retention time of 3.45 min. No additional peak of hydroxylated limonene was detected for any of the samples. Additionally 81D8tr and 81D11tr was tested against limonene, however no conversion was found.

3.4.6 P450 expression in yeast

P450s show a typical absorption maximum at 420 nm. Activity of P450s can be determined by measuring a shift in absorption to 450 nm when CO is bound. P450 transformation into yeast, expression and microsome isolation for CYP71D15 from peppermint was done in Strasbourg and in York and CYP81D11 from Arabidopsis in York. The diluted, reduced microsomes (1:1 in TEG buffer) were analysed by UV/Vis and used for the baseline. The solution was bubbled with CO and analysed again. As a positive control, purified XplA heme was analysed and the typical shift to 450 nm of the CO bound form was observed (Figure 3.19A). There was a higher absorption around 420 nm for CYP71D15 (Figure 3.19B), implying that the P450 was not correctly folded and therefore inactive.
The expression of CYP71D15 from peppermint was repeated in York and additionally CYP81D11 was expressed. During the microsome preparation, surprisingly, the supernatant after the ultracentrifugation of CYP81D11 appeared reddish (the microsomes containing the P450 should be in the pellet) and due to that it was analysed with the UV/Vis spectrophotometer. The supernatant of the ultracentrifugation and the diluted microsomes (1:1 in TEG buffer) with sodium dithionite were used for baseline correction. Then the solution was bubbled with CO and analysed again. However a P450 characteristic spectrum was not obtained (Figure 3.20).

Figure 3.19: CO difference spectrum of the microsomes for XplA heme from Rhodococcus sp. and CYP71D15 from peppermint
A XplA heme (ox = oxidised form, red = reduced form, CO = CO bound form) and B) CYP71D15

Figure 3.20: CO difference spectrum of the microsomes and the supernatant after the ultracentrifugation for CYP81A11 from Arabidopsis and CYP71D15 from peppermint
A CYP81A11 and B CYP71D15
There was no peak at 450 nm for both P450 microsomes, but again a higher absorption around 420 nm. The supernatant (after ultracentrifugation) of CYP81D11 showed also an absorption maximum at around 420 nm.

An SDS PAGE analysis was performed to detect the P450s (Chapter 2.13.2). Three different samples were tested: the supernatant after the sonication of the cells (1 in Figure 3.21), the supernatant after ultracentrifugation (2) and the microsome solution (3).

![SDS PAGE analysis of yeast microsomes containing CYP81D11 (57 kDa) and CYP71D15 PM2-2 (57 kDa).](image)

1 = supernatant after the sonication of the cells, 2 = supernatant after ultracentrifugation, 3 = microsom solution. Bands in the red corners were sent for Maldi-MS analysis.

Signals of the expected sizes for CYP81D11 (57 kDa) and CYP71D15 (57 kDa) could be detected, but MALDI-MS analysis identified the samples as bovine serum albumin and no P450 sequence was detected (Chapter 2.15.2).

3.5 Discussion

3.5.1 P450s expressed in *Escherichia coli*

The function of CYP81D8 and CYP81D11 from Arabidopsis is not known (Appendix A). Both have been found to be induced in response to biotic and abiotic stress, such as osmotic stress, treatment with hydrogen peroxide, jasmonic acid, salicylic acid, abscisic acid, rose bengal and paraquat. Prof. Johnathan
Napier’s research group studied CYP81D11 in more detail and postulated an important role of this P450 in the plant defence response and that it is induced by cis-jasmone (an activator for genes involved in the pathway of secondary metabolic defence chemicals). The P450s CYP81D8, CYP81D11 and the peppermint CYP71D15 PM2-2 were expressed in E. coli and were found in the insoluble protein fraction. There was no improvement in the solubility when CYP81D8 and CYP81D11 were lacking the hydrophobic membrane anchor region. Different E. coli strains were tested to increase the yield of the expressed P450s. The E. coli Rosetta 2 strain, containing a supplementary plasmid for rare codons, resulted in the highest expression yield for 81D8tr and 81D11tr identified by SDS-PAGE when compared with uninduced Rosetta 2 cells.

Additionally, various media were tested for further optimisation of the P450 expression. No expression was achieved in TB, auto-induction and M9 media. The Arabidopsis 81D8tr and 81D11tr have been successfully expressed only in LB medium. However, no signal was seen in western blot analysis, possibly due to P450 protein folding so that the His-Tag cannot be recognised or through detection limitation.

Oxidative activity of 81D8tr and 81D11tr was tested in resting cell assays towards methyl-tolyl-sulphide, because the sulphoxide product is commercial available. It was not necessary to co-express a corresponding plant reductase, due to the native presence of the bacterial flavodoxin reductase (encoded by FPR and used as protection against oxidative stress), which was shown to support plant P450s, for example the CYP97 carotene hydroxylase. However, no activity was found for Arabidopsis 81D8tr and 81D11tr to methyl-tolyl-sulphide in whole cells, which is possibly caused by limitation of the electron transfer through the E. coli reductase or through inactive P450s. Another reason may that methyl-tolyl-sulphide is not a substrate of CYP81D8. Additionally, the P450s, present in the disrupted cell solution after sonication, were tested with two different reductases (commercially available spinach ferredoxin/ferredoxin reductase and recombinant expressed Arabidopsis ATR1tr, Chapter 4), nevertheless no product was detected. Although the expression of both P450s was shown by SDS PAGE analysis, activity could not been detected.
The peppermint CYP71D15 (limonene-3-hydroxylase) is part of the complex pathway leading to the monoterpene (-)-menthol and has been recombinantly expressed in insect cells, yeast and *E. coli* by Prof. Rodney Croteau’s research group. The CYP71D15 construct PM2-2 (provided by Prof. Rodney Croteau) was expressed in *E. coli* following their publications and tested for the hydroxylation of (-)-(4S)-limonene to (-)-trans-isopiperitenol in resting cell assays as well as in different fractions containing protein. No hydroxylation activity was detected possibly caused by too low expression levels, which gave no signal in SDS PAGE and western blot analysis.

### 3.5.2 P450s expressed in *Saccharomyces cerevisiae*

For yeast expression the two *S. cerevisiae* strains WAT11 and WAT21, co-expressing the Arabidopsis reductase ATR1 and ATR2, respectively were used. Several research groups using these two strains have found that the amount of product is dependent on the P450, e.g. the CYP76B1 (from *Helianthus tuberosus*) worked better with WAT11 in the dealkylation of 7-ethoxycoumarin and CYP88A3 and CYP88A4 with WAT21 in the oxidation of ent-kaurenoic acid.

Additional to the activity assays, P450s can be characterised for their typical carbon monoxide difference spectra by shifting the absorption maximum sodium dithionite reduced P450 from 420 nm to the carbon monoxide bound form at 450 nm.

The three P450s CYP89A9 (from Arabidopsis with unknown function), CYP81D11 and CYP71D15 PM2-2 were expressed in *S. cerevisiae* WAT11 and WAT21 and microsomes containing the membrane associated P450s were isolated. P450s were detected by performing a CO difference spectrum. No maximum was seen for these P450s at 450 nm implying that incorrectly folded P450 had been produced. The reason for this is currently unknown. It was hypothesised that the expression of these P450s as fusion proteins would improve expression and facilitate the formation of active protein.
4 Expression, purification and characterisation of truncated, soluble Arabidopsis Cytochrome P450 reductases

4.1 Introduction

Cytochrome P450 reductases (CPR) are important partner enzymes for P450s because they supply electrons for all the different reactions.\textsuperscript{385,386} They are dependent on a flavin cofactor and NAD(P)H as a source of reducing equivalents. In yeast and animal species, only one single reductase is responsible for providing the necessary electrons for all the P450s.\textsuperscript{387} Plant genomes have usually more than one reductase (three in Arabidopsis and also in \textit{Helianthus tuberosus}\textsuperscript{388}), possibly due to the higher number of P450 genes in plant (246 P450 genes in Arabidopsis in comparison to 57 P450 genes in humans or three P450 genes in \textit{S. cerevisiae}).\textsuperscript{389} Angiosperm reductases are usually classified into two distinct classes dependent on their N-terminal amino acid sequence. Class I is usually targeted to membrane of the ER whereas class II contains often more than 20\% of the amino acids Serine and Threonine in the N-terminus and is then anchored in different membranes, such as chloroplasts.\textsuperscript{166,390-392}

In the late 1990s, two Arabidopsis cytochrome P450 reductases, named ATR1 (genetic locus: At4g24520) and ATR2 (At4g30210), were characterised after recombinant expression.\textsuperscript{166,391} A third reductases, ATR3 (At3g02280), was found in the Arabidopsis genome sequence\textsuperscript{393} and was shown to be an authentic reductase supporting cinnamate-4-hydroxylase activity \textit{in vitro}.\textsuperscript{394,395} Low sequence identity (33\%) in the first 86 or 106 amino acids between ATR1 and ATR2, respectively, give rise to the suspicion that the reductases are anchored in different membranes such as endoplasmic reticulum (ER) or chloroplast membrane.\textsuperscript{351} The sequence of ATR2 contains two possible start codons, separated by a Serine and Threonine rich sequence, which is thought to act as membrane signal for chloroplast translocation.\textsuperscript{166,351,396} For this reason, ATR2 belongs to class II reductases whereas ATR1 to class I. Interestingly, ATR3 lacks this hydrophobic N-terminus and is located in the cytoplasm as well as in the nucleus, which was demonstrated by fluorescent spectroscopy of transiently
expressed ATR3 fused to GFP (together with the ER-located mRFP-HDEL ER marker) in tobacco leaves as well as in tobacco By-2 cells. Amino acid sequence analyses demonstrated that ATR1 and ATR2 share 64% sequence identity. The ATR2 amino acid sequence shares an identity of 73% with the P450 reductase from *Catharanthus roseus*\(^{391}\), whereas the ATR1 shares 65% to this reductase. The reductase ATR3 shares only 26% and 24% sequence identity with ATR1 and ATR2, respectively, and belongs rather to the subfamily of diflavin reductases with a sequence identity of 42% to the human nitroreductase NR1 (Figure 4.1).\(^{395}\) ATR3 has been shown to interact with CIAPIN1, a mammalian cytokine-induced inhibitor of apoptosis, and therefore thought to take part in the cell division and programmed cell death.\(^{395}\)
Figure 4.1: Phylogenetic tree of eukaryotic cytochrome P450 reductases with per cent bootstrap value (1000 bootstrap) of 30 CPR sequences and 12 ATR3-like sequences according to Varadarajan et al. 2010 (AA: Artemisia annua, At: Arabidopsis thaliana, Ce: Caenorhabditis elegans, Cery: Centaurium erythraea, Chlamy: Chlamydomonas reinhardtii, Cr: Catharanthus roseus, Dm: Drosophila melanogaster, Gm: Glycine max; Hs, Homo sapiens; Ht, Helianthus tuberosus; Mm, Mus musculus; Mt, Medicago truncatula, Os: Oryza sativa, Ostt: Ostreococcus tauri, Pc: Petroselinum crispum, Pm: Pseudotsuga menziesii, Ps: Pisum sativum, Pso: Papaver somniferum, Ptd: Populus balsamifera subsp. Trichocarpa x Populus deltoids, Rn: Rattus norvegicus, Sc: Saccharomyces cerevisiae, Ta: Triticum aestivum, Tc: Taxus chinensis, Vr: Vigna radiate, Vs: Vicia sativa)

ATR1 and ATR2 have conserved binding motifs for FMN, FAD and NADPH, which are shown in Figure 4.2.
Figure 4.2: Amino acid sequence alignment between the CPR from *C. roseus* and the two Arabidopsis reductases ATR1 and ATR2 modified from Mizutani and Otha 1998. 

### FMN-pyrophosphate

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR1</td>
<td>MVKNTTVVYA</td>
</tr>
<tr>
<td>ATR2</td>
<td>MVKNTTVVYA</td>
</tr>
<tr>
<td><em>C. roseus</em></td>
<td>MVKNTTVVYA</td>
</tr>
</tbody>
</table>

### FMN-isoalloxazine-ring

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR1</td>
<td>ADKIDVYSA</td>
</tr>
<tr>
<td>ATR2</td>
<td>ADKIDVYSA</td>
</tr>
<tr>
<td><em>C. roseus</em></td>
<td>ADKIDVYSA</td>
</tr>
</tbody>
</table>

### Cytochrome c

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR1</td>
<td>MKKTG</td>
</tr>
<tr>
<td>ATR2</td>
<td>MKKTG</td>
</tr>
<tr>
<td><em>C. roseus</em></td>
<td>MKKTG</td>
</tr>
</tbody>
</table>

### FAD-pyrophosphate

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR1</td>
<td>KTFDYV</td>
</tr>
<tr>
<td>ATR2</td>
<td>KTFDYV</td>
</tr>
<tr>
<td><em>C. roseus</em></td>
<td>KTFDYV</td>
</tr>
</tbody>
</table>

### FAD-isoalloxazine-ring

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR1</td>
<td>DQTRVR</td>
</tr>
<tr>
<td>ATR2</td>
<td>DQTRVR</td>
</tr>
<tr>
<td><em>C. roseus</em></td>
<td>DQTRVR</td>
</tr>
</tbody>
</table>

### NADPH-ribose, pyrophosphate

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR1</td>
<td>PFDK</td>
</tr>
<tr>
<td>ATR2</td>
<td>PFDK</td>
</tr>
<tr>
<td><em>C. roseus</em></td>
<td>PFDK</td>
</tr>
</tbody>
</table>

### NADPH-nicotianamide

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR1</td>
<td>QMTGS</td>
</tr>
<tr>
<td>ATR2</td>
<td>QMTGS</td>
</tr>
<tr>
<td><em>C. roseus</em></td>
<td>QMTGS</td>
</tr>
</tbody>
</table>

114
The typical absorption maxima of flavoproteins\textsuperscript{397} at 455 and 380 nm, with the 455 nm peak disappearing when reduced with dithionite had been shown by Mizutani and Ohta\textsuperscript{391} for ATR1 and ATR2. The ATR1 and ATR2 proteins were recombinantly expressed in \textit{Saccharomyces cerevisiae}\textsuperscript{166} and \textit{Spodoptera furugiperda}\textsuperscript{391} insect cells and shown to have activity \textit{in vivo} and \textit{in vitro} with cytochrome c and with the Arabidopsis CYP73A5 (cinnamate-4-hydroxylase). Microarray analysis (https://www.genevestigator.com/gv/) showed that ATR1 is expressed constitutively at all developmental stage whereas ATR2 has more fluctuation in its expression pattern and showed the highest expression in young flowers (Figure 4.3). Expression of ATR3 is at least six times lower at all developmental stages than ATR1 or ATR2.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{expression_levels.png}
\caption{Expression levels of the three Arabidopsis cytochrome P450 reductases ATR1, ATR2 and ATR3 during the plant life cycle (from genevestigator, https://www.genevestigator.com/gv/, 15/12/2011)}
\end{figure}
The full length ATR1 contains 692 amino acids with an estimated size of 77 kDa and ATR2 712 amino acids (79 kDa). The molecular weight of both reductases is similar to that of other higher plant CPRs.

4.2 Objectives

The common host for recombinant expression of plant cytochromes P450 as well as for the partnering reductases are yeast cells (S. cerevisiae) or insect cells. Both are eukaryotic systems which provide all the post-translational machinery for the production of proteins associated with eukaryotic membranes. Expression in E. coli, however, has the advantages that it is an easy to use and an inexpensive expression system. Interestingly, Barnes et al. 1991 reported that P450 reactions are supported in E. coli without the additional expression of a specific reductase. The use of E. coli also has disadvantages. It is a prokaryotic expression system and expression can be hindered by different codon bias. Nevertheless, soluble enzyme has been produced in bacteria following deletion of the hydrophobic N-terminus encoding the membrane anchor region.

In this chapter, the two membrane-associated reductases ATR1 and ATR2 were cloned from Arabidopsis and modified at the N-terminus to increase solubility. The truncated versions (ATR1tr and ATR2tr) were overexpressed in E. coli, purified and activities compared to find the most effective reductases for expression in E. coli. Furthermore, the biochemistry of both reductases was studied.

4.3 Results

4.3.1 Primer design for the N-terminal membrane anchor truncated ATR1 and ATR2

The two Arabidopsis P450 reductases ATR1 (77 kDa) and ATR2 (79 kDa) have both a hydrophobic N-terminal region in the amino acid sequence, which allows association with a membrane. The amino acid sequence of both reductases was analysed using the software TMHMM and SignalP3.0 (Chapter 2.8.1) (Figure 4.4) and the hydrophobic N-termini removed (46 and 73 aa were deleted for ATR1 and ATR2, respectively).
A amino acid sequence of ATR1

MTSALYASDLFKLSIMGTDSLSDDVLVIATTSLALVGAVFVLWKKTADRSQELKPLMIPKSLMADDDDLGSKTRVSIFGFTQGTTAGFAKALSEERIKYAEKAKVVIDLDAYADDDQVEKLKKEFETAFFCVAYGVGDGEPTDNAARFYKHFMKEDDIKQLLQAYGVFALGNRQYEHFNKIGIVLDEELECKGAKRLIEVGLGDDQSEIDDEFNAWKEJLMS

ELDQLLKKDEDDKSVATFYTAPEIYKVRTHDPRFTQTQKSMESNVANGNTTIDIHHPFRCVDVAVQHELHTESDRCIHLEFDSIRGITYETGDHVGYAENHVIEEAGKLLLHSLDLVIHADKEGSRSLAESAPFFPGFCTLGTLALRAYAADDLNFFRKSALVQALAAAYATEPSEAELKHLLTSPDGKDEYSQWIVASQRSLLEVMAAFFFSP(allGVFFAAIAIPRLQFRYY

SISSPRLAPSRVHVTSDLVGPTPTGRHKGVCSTWKNVFAKSEHECSGAIFIRASNFKLQPSSTPIVMGFGTGALGPFPGLQERMALKEDGEGLGLSLLFQGCRMRQMDFIYEDELNNFVDQGVISELMAFSREGAQKEYQVKMMEKAQQVWDLIKEEGYLYVCDAKGMARDVHRTLHITVQCEEQGVSSEEAIKVLQTEGRYLRDVW

B TMHMM result

![TMHMM posterior probabilities for ATR1](image)

C SignalP result

![SignalP-HMM prediction (euk model)](image)

Figure 4.4: Amino acid sequence analysis of ATR1 N-terminus.

A full length amino acid sequence of ATR1, B TMHMM blot for analyzing the amino acid sequence of ATR1, which indicated a membrane region for around the first 50 amino acids, C result of SignalP 3.0 analysis (cleavage prob = probability of cleavage site, n-region prob = probability of N-terminus of the signal peptide, h-region prob = probability of hydrophobic region of signal peptide, c-region prob = probability of C-terminus of signal peptide)
4.3.2 Cloning of ATR1tr and ATR2tr

Genes encoding Arabidopsis reductases ATR1 and ATR2 were cloned following RNA extraction and reverse transcription into cDNA. The N-terminal truncated cDNAs of ATR1tr and ATR2tr were amplified using PCR (Chapter 2.8.2) and the truncated reductases were cloned into the LIC-vector using Ligation Independent Cloning (Chapter 1.1).

4.3.3 Expression of ATR1tr and ATR2tr in Escherichia coli

ATR1tr and ATR2tr were expressed at 20 °C in the two different E. coli expression strains Rosetta 2 (DE3) and Rosetta gami 2 (DE3) (conditions see Chapter 2.10.1).

The induced culture showed inhibited cell growth in comparison to the uninduced culture (Figure 4.5).

![Figure 4.5: Cell growth (OD600) after induction of ATR1tr in E. coli Rosetta 2 grown in LB medium](image)

The expression of the 72 kDa ATR1tr in Rosetta 2 (DE3) cells was visible by SDS-PAGE (Figure 4.6A) and no expression could be detected in the uninduced control (Figure 4.6B). No overexpression of either reductases could be achieved when using strain Rosetta gami 2 (DE3), so Rosetta 2 (DE3) was used for the following experiments. Low amounts of ATR1tr could be expressed in Rosetta 2 (DE3) cells when M9 minimal media was used (Figure 4.6C and D).
ATR1tr was purified with Ni-resin material (Chapter 2.10.4). Purification was monitored using SDS PAGE (Figure 4.7) and showed that ATR1tr was partly soluble, which was purified using Ni-resin material. ATR1tr precipitated during dialysis and was again insoluble. To overcome these problems, different solubilisation buffers were tested (Chapter 2.10.2).
**Figure 4.7:** SDS PAGE analysis of ATR1tr (72kDa) during Ni-affinity purification

M marker, 1 whole cells prior sonication, 2 soluble protein fraction, 3 insoluble protein fraction, 4 eluted protein, 5 dialysed protein and 6 precipitated protein fraction after dialysis

### 4.3.4 Solubilisation of overexpressed ATR1tr

Thirty different buffers (Chapter 2.10.2) were tested to increase the solubility of the overexpressed ATR1tr and ten out of 30 buffers increased the concentration of soluble protein (Figure 4.8).

**Figure 4.8:** Protein concentration of the soluble total protein after overexpression in *E. coli* Rosetta 2 (DE3) using different solubilisation buffers 0-30
Protein fractions from all 30 buffers were analysed using SDS PAGE. The strongest signal for ATR1tr was obtained when using the buffers 1, 2, 6, 16, 17, 18, 20, 24, 25 and 28 (Figure 4.9). Western blot probing using anti-His antibodies confirmed these results (Figure 4.10). No signal was detected for the negative control containing the empty plasmid when resuspended in buffer 0 or solubilised in buffer 18 (lane 0c and 18c, respectively in Figure 4.10).

**Figure 4.9: SDS PAGE analysis of the total soluble protein using solubilisation buffer 0-30. Red box show the overexpressed ATR1tr (72 kDa)**

**Figure 4.10: Analysis of ATR1tr after sonication in different solubilisation buffers**

A Western blot analysis and B SDS PAGE of soluble protein fraction. Red boxes show the overexpressed ATR1tr. (Lane 0c = negative control in buffer 0, 18c = negative control in buffer 18, 0-28 = overexpressed ATR1tr in different solubilisation buffers)

4.3.5 Purification of ATR1tr and ATR2tr in solubilisation buffer 18A

ATR1tr and ATR2tr were overexpressed under the same conditions described in Chapter 2.10.1 and using buffer 18 to assist in the solubilisation of protein. However, buffer 18 contains EDTA, making it unsuitable for the purification with His-Select Nickel Affinity Gel column. EDTA can scavenge metal ions and
therefore complexes the nickel ions so that the His-tagged protein cannot bind. Therefore, buffer 18 without EDTA (buffer 18A) was used for the purification of ATR1tr and ATR2tr (Chapter 2.10.5).

A SDS PAGE analysis was performed after the purification of both reductases and ATR1tr and ATR2tr were detected (red box in gel A and B, Figure 4.11). The empty vector transformed into Rosetta 2 (DE3) was used as negative control (Figure 4.11C).

Figure 4.11: SDS PAGE Analysis of the purification of A) ATR1tr (72 kDa), B) ATR2tr (71 kDa) and C) negative control (empty vector expressed in *E. coli* Rosetta 2 (DE3))
4.3.6 Characterisation of Arabidopsis P450 reductase ATR2tr

4.3.6.1 Activity assay with cytochrome c

The activity of ATR1tr and ATR2tr with cytochrome c was tested for the two different co-factors NADH and NADPH following the protocol described in (Chapter 2.15.1.2) to analyse the impact on the activity. Both Arabidopsis reductases displayed greater activity with NADPH over NADH as cofactor (Figure 4.12).

![Figure 4.12: ATR1tr and ATR2tr activity dependence on the cofactor NADH and NADPH](image)

Assay conditions: 1mg/ml reductase, 50 μM cytochrome c, 110 μM NAD(P)H in 300 mM potassium phosphate buffer (pH 7.7). The error bars represent the mean of three independent replicas ± standard deviation.
CHAPTER 4 – ARABIDOPSIS REDUCTASES

ATR1tr and ATR2tr activity was compared with two commercial available reductases (ferredoxin reductase from spinach and cytochrome c reductase from porcine heart). Both, ATR1tr and ATR2tr, showed higher activity in comparison to the positive control reductases (Figure 4.13). ATR2tr was diluted from 1 mg/ml to 0.01 mg/ml to get a linear rate. The activity of cytochrome c reductase was not significant due to the fact that it showed a higher activity with the cofactor NADH than with NADPH.

![Graph showing cytochrome c conversion by different reductases](image)

**Figure 4.13: Cytochrome c conversion by different reductases in presence of the cofactor NADPH**

Assay conditions: reductase (amount, see legend), 50 μM cytochrome c in 300 mM potassium phosphate buffer (pH 7.7). The reaction was started after one minute by adding 110 μM NADPH (ferr red = ferredoxin reductases from spinach, cytC red = cytochrome c reductase from porcine heart).
ATR2tr was with a conversion of 1500 μM cytochrome c/mg ATR2tr 20 times more active than ATR1tr (68 μM cytochrome c/min/mg ATR1tr, Figure 4.14). Porcine cytochrome c reductase and spinach ferredoxine reductase was used as positive control and showed a conversion of 0.5 μM and 7.5 μM cytochrome c/min/mg reductase, respectively. The low activity of cytochrome c reductase is caused by the cofactor NADPH. Higher values were detected by using the cofactor NADH.

Figure 4.14: Cytochrome c conversion by different reductases in presence of the cofactor NADPH
ATR1tr = N-terminal truncated Arabidopsis reductase ATR1, ATR2tr = N-terminal truncated Arabidopsis reductase ATR2, cytochrome C red = cytochrome c reductase from procine heart, ferredoxin red = ferredoxin reductase from spinach
### 4.3.6.2 Temperature and pH optima of ATR1tr and ATR2tr

Temperature and pH optima for ATR1tr and ATR2tr were performed as described in Chapter 2.15.1.3. The temperature optimum of ATR1tr and ATR2tr was determined to be 40 °C (Figure 4.15).

![Temperature optimum for ATR1tr (1 mg/ml) and ATR2tr (0.01 mg/ml)](image)

Figure 4.15: Temperature optimum for ATR1tr (1 mg/ml) and ATR2tr (0.01 mg/ml)
The error bars represent the mean of three independent replicas ± standard deviation

The pH optima of both Arabidopsis reductases was analysed in Britton Robinson buffer (Chapter 2.15.1.3) covering a pH range of pH 3.4 to 11.0. The extinction coefficient of cytochrome c ($\varepsilon_{(550\text{nm})} = 21 \text{ mM}^{-1}\text{cm}^{-1}$) was used to calculate the activity.

The pH optima of ATR1tr (100 μg/ml) and ATR2tr (1 μg/ml) were detected in 300 mM Britton Robinson buffer (Chapter 2.15.1.3). The optima of ATR1tr and ATR2tr were found to be pH 7.0 and pH 8.0, respectively (Figure 4.16).
The reductase activity of ATR2tr was also tested in citrate and phosphate buffer, covering a pH range of 4.0 to 7.6 (Chapter 2.15.1.3). ATR2tr showed approximately 10% more activity in the citrate and phosphate buffer when compared to Britton Robinson buffer (Figure 4.17).

Figure 4.16: Activity of 100 μg/ml ATR1tr and 1 μg/ml ATR2tr in 300 mM Britton Robinson buffer (1:1:1 mixture of boric acid, phosphoric acid and acetic acid) at different pH values
The error bars represent the mean of three independent replicas ± standard deviation.

Figure 4.17: Activity of 1 μg/ml ATR2tr in diverse buffers covering different pH values
The error bars represent the mean of three independent replicas ± standard deviation.
The activity of cytochrome c as well as the stability of the cofactor NADPH, both part of the reductase activity assay, are also pH dependent and can therefore change the pH optima of the two reductases.

Cytochrome c is changing the conformational state dependent on the pH and has an isoelectric point (number of positive and negative charges of the protein are equal) around pH 10.\textsuperscript{400–403} NADPH degrades at low pH values (Figure 4.18) and so fewer electrons are available for the reaction. Therefore, the reductases activity at a pH lower than pH 6.0 was reduced through the instability of NADPH.

![Stability of the cofactor NADPH at different pH values](image)

**Figure 4.18: Stability of the cofactor NADPH at different pH values**
Assay was performed using 50 μM cytochrome c and 110 μM NADPH in 300 mM potassium phosphate buffer (pH 7.7). The error bars represent the mean of three independent replicas ± standard deviation.

### 4.3.6.3 Kinetic studies of the reductases ATR1tr and ATR2tr

The kinetic data for cytochrome c reduction were accomplished following the published protocol from Hull and Calenza\textsuperscript{351} (Chapter 2.15.1.4). \(V_{\text{max}}\) and \(K_M\) values were determined using Michaelis-Menten kinetics (software: GraFit) (Table 4.1). ATR2tr followed the common Michaelis-Menten kinetic, whereas ATR1tr was inhibited by the substrate cytochrome c for concentrations higher than 100 μM (Figure 4.19). The sample with the highest substrate concentration (150 μM cytochrome c) was ignored for the calculation of the kinetic data for ATR1tr by Michaelis-Menten (Figure 4.19A).
Table 4.1: Kinetic data for ATR1tr and ATR2tr (calculated by the program GraFit)

<table>
<thead>
<tr>
<th>reductase</th>
<th>$V_{\text{max}}$ (μM/min)</th>
<th>$K_M$ (μM)</th>
<th>inhibition constant</th>
<th>$k_{\text{cat}}$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR1tr (Michaelis-Menten)</td>
<td>51 ± 5</td>
<td>23 ± 6</td>
<td></td>
<td>0.1 ± 0.02</td>
</tr>
<tr>
<td>ATR1tr (substrate inhibition)</td>
<td>261 ± 90</td>
<td>170 ± 68</td>
<td>21 ± 9</td>
<td>0.3 ± 0.1</td>
</tr>
<tr>
<td>ATR2tr (Michaelis-Menten)</td>
<td>38 ± 2</td>
<td>16 ± 3</td>
<td></td>
<td>5.0 ± 0.6</td>
</tr>
</tbody>
</table>

Figure 4.19: Michaelis-Menten blots by GraFit for A) ATR1tr by Michaelis-Menten equation and B) ATR1tr by equation for substrate inhibition and C) ATR2tr by Michaelis-Menten equation
To illustrate the accuracy of the kinetic data, the direct linear blot from Eisenthal and Cornish-Bowden was used. Here the median of all samples is used to estimate the $K_M$ and $V_{max}$. As shown in Figure 4.20A the intersection for ATR1tr did not match as good as for ATR2tr (Figure 4.20).

![Figure 4.20: Cornish-Bowden blot for A) ATR1tr and B) ATR2tr](image)

The different lines represent substrate concentrations between 4 μM and 150 μM cytochrome c.

The specific activity for ATR1tr considering substrate inhibition was 50.5 μM min$^{-1}$ mg$^{-1}$ and for ATR2tr 3770 μM min$^{-1}$ mg$^{-1}$.
4.3.6.4 Spectrophotometric characterisation of ATR2tr

P450 reductases have a typical absorption spectrum with a maximum at 375 and 450 nm. Dialysed ATR2tr (5.1 mg/ml in 10 mM potassium phosphate buffer, pH 7.5) was characterised for the presence of flavin by monitoring the absorption between 300 and 600 nm (Figure 4.21).

![Figure 4.21: UV/Vis spectra of purified recombinant ATR2tr and negative control (expressed and purified empty vector)](image)

The diagram illustrated a similar picture to published data with two absorption maxima, characteristic for flavins\(^{391}\) at 383 and 458 nm. FMN and FAD alone have absorption maxima at around 370 and 452 nm (Figure 4.22).

![Figure 4.22: UV/Vis spectra of 10 \(\mu\)M FMN and 10 \(\mu\)M FAD in 50 mM potassium phosphate buffer pH 7.5](image)
4.3.6.5 Activity test of purified ATR2tr in buffer A and B

As previously described, ATR2tr was purified using a column containing Ni-resin material, eluted with buffer 18A (100 mM Tris, 100 mM KCl and 1% Triton X-100, pH 8.2) containing 500 mM imidazole and dialysed in buffer A (50 mM potassium phosphate buffer pH 7.5). Hull and Calenza (2000) suggested another dialysis buffer of 30 mM potassium phosphate pH 7.8, 20% glycerol, 0.1 mM EDTA, 2.0 μM FMN (called buffer B) stabilising the Arabidopsis reductase ATR2.351

The reductase activity was tested using a cytochrome c assay following the protocol from Hull and Calenza 2000 (Chapter 2.15.1.2).351 ATR2tr activity was inhibited in the presence of 500 mM imidazole in 100 mM Tris, 100 mM KCl and 1% Triton X-100, pH 8.2 (Figure 4.23). The highest activity for ATR2tr was obtained after the dialysis in buffer B.

![Figure 4.23: Michaelis-Menten-diagram of purified ATR2tr](image)

In buffer 18A (100 mM Tris, 100 mM KCl and 1% Triton X-100, pH 8.2) containing 500 mM imidazole (blue rhombi), buffer A (50 mM potassium phosphate buffer pH 7.5, green triangles) and buffer B (30 mM potassium phosphate pH 7.8, 20% glycerol, 0.1 mM EDTA, 2.0 μM FMN, red squares). The error bars represent the mean of three independent replicas ± standard deviation.
4.3.6.6 Stability test of ATR2tr in different buffers

The reductase ATR2tr was analysed after dialysis in two different buffers: a general potassium phosphate buffer pH 7.5 (buffer A) and a potassium phosphate buffer pH 7.8 containing FMN and glycerol (buffer B). The samples were stored at four different temperatures (−80 °C, −20 °C, 4 °C and 21 °C) and the activity of ATR2tr was measured using a cytochrome c assay (Chapter 2.15.1.4). ATR2tr was stable at 21 °C for at least two days and was inactive after less than six days independent from the storage buffer (Figure 4.24). A 75% and 50% decrease of the reductase activity was observed when stored at 4 °C in buffer A and B, respectively. ATR2tr was less stable in buffer A and lost activity steadily over time (Figure 4.24A). When ATR2tr was kept at −20 °C for 40 days, the activity was reduced to 30%, whereas storage at −80 °C meant that the activity was only reduced to 50% after the same period of time. A 30% loss in activity of ATR2tr was observed upon freezing at −20 °C or −80 °C in buffer B; however, no further reduction in activity was found over time. In summary, ATR2tr was more stable in buffer B due to the presence of FMN and glycerol.

Figure 4.24: Stability of ATR2tr over time in A) buffer A (50 mM potassium phosphate buffer pH 7.5) and B) buffer B (30 mM potassium phosphate pH 7.8, 20 % glycerol, 0.1 mM EDTA, 2.0 μM FMN) The error bars represent the mean of three independent replicas ± standard deviation. Sample at time 0 days was not frozen and activity measured at 21 °C.
4.4 Discussion

The two Arabidopsis reductases ATR1 and ATR2 are necessary partner enzymes for all the different P450 reactions in Arabidopsis and supply electrons in single steps from the co-factor NADPH to the heme iron in the catalytic centre of the P450 enzyme. Both reductases contain a membrane signal at the N-terminus and although active full length forms have been expressed previously in the eukaryotic hosts yeast\textsuperscript{166} and insect cells\textsuperscript{391}, the proteins were predominantly insoluble. Activity assays established high stability for ATR1 while activity for ATR2 rapidly decreased when expressed in a eukaryotic host.\textsuperscript{350} Therefore, to increase the amount of soluble protein the hydrophobic N-termini of both enzymes were removed in this project and the two truncated reductases ATR1\textsubscript{tr} and ATR2\textsubscript{tr} (lacking 46 and 73 amino acids, respectively) expressed in \textit{E. coli}.

Different \textit{E. coli} expression strains were tested and the \textit{E. coli} Rosetta 2 (DE3) strain was found to produce the highest yield for both truncated reductases. The recombinant, expressed enzymes were partly insoluble, although the N-terminus was deleted, and so 30 different solubilisation buffers were tested with ATR1\textsubscript{tr} to increase the solubility. The highest level of soluble ATR1\textsubscript{tr}, as confirmed by SDS PAGE and western blot analysis, was seen in buffer 18 containing 100 mM Tris, 100 mM KCl, 2 mM EDTA and 1% Triton X-100 (pH 8.2).

Both ATR1\textsubscript{tr} and ATR2\textsubscript{tr} were subsequently purified in buffer 18A (buffer 18 without the addition of EDTA). The absence of EDTA in buffer 18A did not affect the ability to solubilise ATR1\textsubscript{tr} and ATR2\textsubscript{tr}, as this characteristic was due to the high concentration (1\%) of the detergent Triton X-100. Triton X-100, a non-ionic detergent used to solubilise membrane proteins.\textsuperscript{405}

The temperature optima for both reductases was found to be at 40 °C and pH optima of ATR1\textsubscript{tr} and ATR2\textsubscript{tr} were pH 7.0 and pH 8.0, respectively. This pH optimum was also found to be similar for other plant reductases, such as the CPR from \textit{Coleus blumei} with pH 7.5.\textsuperscript{337}

Activity studies on the two N-terminal truncated Arabidopsis reductases showed that both were active with cytochrome c and NADH or NADPH, whereas the activity with NADPH as cofactor was 75 and 830 times higher for ATR1\textsubscript{tr} and ATR2\textsubscript{tr}, respectively. That NADPH is the more effective cofactor is not always the case, cytochrome c activity was better supported by NADH for example for
the CPR from cotton and the CPR from pea. Recombinant expressed ATR2tr was 20-fold more active than ATR1tr, or commercial available ferredoxin reductases. Kinetic values for ATR1tr and ATR2tr were determined with \( V_{\text{max}} \) 51 and 38 \( \mu \text{M/min} \) and \( K_M \) 23 and 16 \( \mu \text{M} \), respectively, using Michaelis-Menten kinetics. ATR1tr showed substrate inhibition, therefore, the interpretation of the data should be treated with caution. The kinetic data of ATR2tr were similar to published data of native ATR2 expressed in \textit{S. cerevisiae} (\( K_M \) 15 \( \mu \text{M} \)) and of solubilised ATR2mod from the membrane fraction expressed in \textit{E. coli} (\( K_M \) 15 \( \mu \text{M} \) and 9 \( \text{U/mg} \)). The latter authors deleted only 45 amino acids of the predicted chloroplast signal, whereas in this project 73 amino acids (the whole hydrophobic membrane anchor) were removed. This shows that the N-terminus is not part of the catalytic centre, which was also demonstrated for the reductase of \textit{S. cerevisiae}. The native reductase and a truncated soluble reductase (22 amino acids deleted at the N-terminus) of \textit{S. cerevisiae} resulted in similar kinetic values for the activity with cytochrome c and CYP61 (forming ergosta-5,7-dienol to ergosterol). Soluble ATR2tr (\( K_M \) 16 \( \mu \text{M} \)) has a 1.4 to 14 lower affinity to cytochrome c compared to the \( K_M \) values of other soluble and truncated eukaryotic P450 reductases recombinantly expressed in \textit{E. coli} (1.1 \( \mu \text{M} \) for \textit{S. cerevisiae} CPR and 12 \( \mu \text{M} \) for \textit{Phanerochaete chrysosporium} CPR). Recombinantly expressed and then solubilised full length CPRs showed also a stronger affinity of cytochrome c, such as 5 \( \mu \text{M} \) for house fly CPR. Overexpressed and purified plant CPRs gave with 7 \( \mu \text{M} \) (\textit{C. roseus}) two times higher affinity than ATR2tr (16 \( \mu \text{M} \), this work).

ATR2 has two possible start codons, so two possible proteins can be produced. ATR2-1 (712 aa) and ATR2-2 (702 aa) were expressed in \textit{S. cerevisiae} reported by Urban \textit{et al.} 1997. The specific microsomes activity with cytochrome c was 2.8 \( \mu \text{mol} \cdot \text{l}^{-1} \cdot \text{min}^{-1} \cdot \text{mg}^{-1} \) and 0.3 \( \mu \text{mol} \cdot \text{l}^{-1} \cdot \text{min}^{-1} \cdot \text{mg}^{-1} \) for ATR2-1 and ATR2-2, respectively. ATR2tr expressed in \textit{E. coli} (this project) showed with 3770 \( \mu \text{mol} \cdot \text{l}^{-1} \cdot \text{min}^{-1} \cdot \text{mg}^{-1} \) more than 1000 higher specific activity than native ATR2-1 (2.8 \( \mu \text{mol} \cdot \text{l}^{-1} \cdot \text{min}^{-1} \cdot \text{mg}^{-1} \)) or ATR2-2 (ten amino acids on the N-terminus deleted, 0.3 \( \mu \text{mol} \cdot \text{l}^{-1} \cdot \text{min}^{-1} \cdot \text{mg}^{-1} \)) expressed in yeast microsomes. Albeit, characterisation studies of ATR1 and ATR2 resulted in similar specific activity for the conversion
of cinnamic acid to coumaric acid when tested with the Arabidopsis P450 cinnamate-4-hydroxylase (CYP73A5), possibly due to the limited velocity of the P450.\textsuperscript{166,391} ATR2\textsubscript{tr} showed also a higher specific activity when compared with other plant reductase, e.g. reductase of \textit{Petunia hybrida} (30-60 \textmu mol\textperiodcentered l\textsuperscript{-1}\textperiodcentered min\textsuperscript{-1}\textperiodcentered mg\textsuperscript{-1})\textsuperscript{412} and \textit{Mentha spicata} (43 \textmu mol\textperiodcentered l\textsuperscript{-1}\textperiodcentered min\textsuperscript{-1}\textperiodcentered mg\textsuperscript{-1})\textsuperscript{413}. Moreover, a lower specific activity was reported for P450 reductases from other organisms, for example 2.85 \textmu mol\textperiodcentered l\textsuperscript{-1}\textperiodcentered min\textsuperscript{-1}\textperiodcentered mg\textsuperscript{-1} for human reductases recombinant expressed in insect cells\textsuperscript{314} or 10.7 \textmu mol\textperiodcentered l\textsuperscript{-1}\textperiodcentered min\textsuperscript{-1}\textperiodcentered mg\textsuperscript{-1} \textit{Candida maltose} P450 reductase in microsomes after expression in \textit{S. cerevisiae}\textsuperscript{415}.

The $K_M$ value for ATR2\textsubscript{tr} found in this work was 16 \textmu M and a similar value was found when expressed in yeast\textsuperscript{350}. The turnover number $k_{cat}$ (the maximum number of enzymatic reactions per second) was found to be around three times lower for ATR2\textsubscript{tr} (5 s\textsuperscript{-1}) than for native ATR2 expressed in yeast (14 sec\textsuperscript{-1})\textsuperscript{350}. ATR2\textsubscript{tr}, after dialysis in buffer B, exhibited more activity than when assayed in general potassium phosphate buffer A due to its low affinity for the coenzyme FMN\textsuperscript{350}. ATR2\textsubscript{tr} expressed in \textit{E. coli} and stored in buffer B at \textdegree 21 C or \textdegree 80 C kept 70\% of the original reductases activity using cytochrome c over six weeks (Chapter 4.3.6.6).

In conclusion, the two recombinant expressed Arabidopsis reductases ATR1\textsubscript{tr} and ATR2\textsubscript{tr} were soluble and active after purification. ATR1\textsubscript{tr} showed substrate inhibition with cytochrome c and was 200 times less active than ATR2\textsubscript{tr}, which performed typical Michaelis-Menten kinetics. The kinetic data of ATR2\textsubscript{tr} were similar to published data, whereas the specific activity was higher and it was found to be more stable than reported in the literature. Therefore, ATR2\textsubscript{tr} was selected for the creation of the platform technology containing plant P450s and an Arabidopsis reductase (Chapter 5).
5 Development of the ATR2tr-LIC platform

5.1 Introduction

For industry, *E. coli* is generally the preferred heterologous host for recombinant protein expression and this robust bacterial expression system would be desirable for the characterization of Arabidopsis P450s. Plant P450s have been expressed as membrane associated proteins in *E. coli* (more information in Chapter 3). Unfortunately, there is no evidence in the literature that plant P450s can be soluble and actively expressed in *E. coli*. The difficulties in expression of plant P450s in bacteria are caused by the hydrophobic membrane anchor, the complex protein structure and different codon bias between plant and bacteria. Due to this challenge, plant P450s have been mainly expressed in mammalian\(^{36,416}\), insect\(^{191,416}\) or yeast\(^{166,328}\) cells, but it also has been shown that it is possible to engineer plant P450 systems for expression in bacteria as membrane associated proteins\(^{299}\).

As a reductase is necessary for the P450 reaction it would be useful to fuse a P450 of choice to a reductase to create a self-sufficient system. In 1995, the first artificial plant P450 fusion enzyme was recombinantly expressed in *E. coli* and activity detected in membrane preparations.\(^{299}\)

The following plant cytochromes P450 are described in this Chapter for the development of the platform technology where plant P450s are fused to Arabidopsis reductase ATR2tr: isoflavone synthase I (IFS), cinnamate-4-hydroxylase (CYP73A5), N-demethylase (CYP82E4) and CYP81D8 with unknown function.

The native isoflavone synthase I (IFS) from *Glycine max*, which was already used for the artificial fusion of the plant reductases CPR from *Catharanthus roseus* by Prof. Koffas’ group\(^{306}\) was chosen. The enzyme IFS catalyses the two aryl migrations (from liquiritigenin to daidzein and from naringenin to genistein) in the phenylpropanoid pathway (Figure 5.1). The isoflavones daidzein and genistein belong to the class of phytoestrogens and have an antioxidant effect, which protects the DNA in the cells from oxidative damage.
Figure 5.1: Phenylpropanoid pathway showing the key role of the two P450s IFS (CYP93C) and C4H (CYP73A5) used in this project

PAL = phenylalanine ammonia lyase, C4H = cinnamate-4-hydroxylase, 4CL = 4-coumarate-CoA-chalcone isomerase, CHS = chalcone synthase, CHR = chalcone reductase, CHI = chalcone isomerase, IFS: isoflavone synthase

The construct was kindly provided by Prof. Mattheos Koffas (State University of New York, Buffalo, USA) as the artificial fusion IFS-CPR E306 (IFS = isoflavone synthase I from G. max fused to the P450 reductase from C. roseus). Modifications of the anchor region at the N-terminus of the P450 IFS gene, as well as the linker between the P450 and the reductase, were altered to increase the isoflavone synthase activity. The highest activity towards naringenin was seen
when an N-terminal modified IFS was used. Therefore, the first six N-terminal residues were replaced by the synthetic mammalian peptide (amino acid sequence: MALLLAVF; \(\varepsilon\)) and then fused to a truncated \(C.\ roseus\) reductase through the linker \(\lambda\), encoding glycine-serine-threonine (Figure 5.2). \(^{306}\)

**Figure 5.2: Arrangement of IFS-CPR fusion construct**
IFS = isoflavone synthase I, \(\varepsilon\) = synthetic sequence encoding mammalian peptide, \(\lambda\) = linker, encoding glycine-serine-threonine, CPR = \(Catharanthus\ roseus\) reductases (Leonard and Koffas \(^{306}\))

IFS (GenBank accession number AF195798) belongs to the CYP93C subfamily and is closely related to CYP93C1 (97% amino acids sequence, AF022462) \(^{417} 418\). The native \(IFS_{nat}\) as well as the codon optimised \(IFS_{opt}\) (GeneArt) for expression in \(E.\ coli\) were both used in these studies.

The cinnamate-4-hydroxylase (CYP73A5) was selected, because it is a well studied P450 from Arabidopsis \(^{419}\) and is involved in the phenylpropanoid pathway (Figure 5.1), hydroxylating \(trans\)-cinnamic acid to \(p\)-coumaric acid. It was characterised in 1997 after recombinant expression in insect cells. \(^{191}\)

The CYP82E4 from \(Nicotiana\ tabacum\) was chosen due to its N-demethylase activity of nicotine (Figure 5.3), a desired reaction for the pharmaceutical industry. Tertiary N-methylamine products from alkaloids are attractive for the pharmaceutical industry, such as alkaloids from the opiate family, for example the analgesics morphine \(^{420}\) and codeine \(^{421}\), or tropanes, for example cocaine \(^{422}\) and the central nerve system stimulating atropine \(^{423}\). They are generally produced with different chemical methods that often require toxic reagents such as cyanogen bromide \(^{422}\) or chloroformate esters \(^{424}\).

**Figure 5.3: N-demethylation of nicotine to nornicotine by CYP82E4 from Nicotiana tabacum**
The alkaloid nicotine is the main representative in tobacco products. Nornicotine can be reduced chemically or when incinerated to nitroso products, such as N’-nitroso-nornicotine, which are toxic and mostly carcinogenic.

The function of CYP81D8 is not known yet. Nevertheless, this P450 gene was found to be more than 30 fold upregulated in Arabidopsis after treatment with the explosive 2,4,6-trinitrotoluene (TNT) in a microarray experiment (Chapter 3.1, Table 3.2). TNT is an artificial substance, which is highly toxic. It persists in the environment and so far no enzyme has been identified able to mineralise this compound. Due to the fact that CYP81D8 is responding to TNT, it may be that this P450 is also involved in the transformation of other organic xenobiotics.

5.2 Objectives

Plant cytochromes P450 are known to catalyse a diverse range of chemistries and are attractive targets for oxidative biocatalysis and detoxification of environmental pollutants. However, as plant P450s are membrane associated, recombinant expression of soluble protein is problematic and as a result very few of these enzymes have been characterised. A strategy to obtain soluble protein when expressed in E. coli is the removal of the N-terminal hydrophobic membrane anchor region at a molecular level (Chapter 2.8.1). Fusing the N-terminal truncated heme domain of plant P450s with an appropriate N-terminal truncated reductase could provide the basis for developing robust, soluble redox-self-sufficient plant enzyme systems for characterisation studies.

This chapter describes the development of a new technology platform using a Ligation Independent Cloning (LIC) vector (Chapter 2.9), in which the truncated version of the Arabidopsis P450 reductase ATR2tr was integrated. This novel ATR2tr-vector was tested with a selection of plant P450s catalysing different P450 reaction types. The technology could be used in a high throughput screen of Arabidopsis P450s with the aim of determining plant P450 functions and engineering biocatalysts for industrial applications.
5.3 Methods

5.3.1.1 Developing the platform technology of plant P450-ATR2tr fusions

There are three cytochromes P450 reductases encoded in the Arabidopsis genome. A truncated version of the Arabidopsis reductase ATR2 was selected for developing the ATR2tr-LIC vector because studies (Chapter 4) had found that recombinant ATR2tr had approximately 20 times more activity than ATR1tr. ATR3 was not tested, but microarray analysis suggests that it has lower expression (Figure 4.3).

The Arabidopsis reductase ATR2tr was synthesised by GeneArt (Regensburg, Germany) to remove the four native BseRI restriction sites and to codon optimise the sequence for expression in *E. coli*. ATR2tr was cloned into the LIC-vector using traditional cloning methods (described Chapter 2.8) with BseRI (N-terminus) and NdeI (C-terminus) restriction sites.

The linker was inserted at the N-terminal sequence of ATR2tr by PCR. The forward primer (Lic_BseRI_F, Table 5.1) was designed to integrate the lic-linker and two primers for the lam-linker (lam_BseRI_F1 and lam_BseRI_F2, Table 5.1), due to its length. The reverse primer (Atr2_NdeI_Lic_R) was designed with a NdeI restriction site and lamATR2tr and licATR2tr was amplified by PCR using Phusion polymerase (Chapter 2.8.2).

<table>
<thead>
<tr>
<th>primer name</th>
<th>sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lic_BseRI_F</td>
<td>AGAGAGGCGGCCTCCTCTCCTCAAGGTAGCGGTAATAGCAACGTGTTG</td>
</tr>
<tr>
<td>lam_BseRI_F1</td>
<td>CTCTCCAGGATCGGGTGTTAGCGGTAATAGCAACGTGTTG</td>
</tr>
<tr>
<td>lam_BseRI_F2</td>
<td>AGAGAGGCGCCTCCTCTCCAGGATCGGGTGGTAGCAACGTGTTG</td>
</tr>
<tr>
<td>Atr2_NdeI_Lic_R</td>
<td>GAGAGACATATGGTTACAAACATCACCAGATAACG</td>
</tr>
</tbody>
</table>
The ATR2tr inserts were purified (Chapter 2.6) and digested sequentially. Therefore, 13 μg insert DNA, 15 μl NEB4 buffer (New England Biolabs), 5 U NdeI (New England Biolabs) and water to a final volume of 150 μl were incubated overnight at 37 °C. Then a further 5 μl NEB4 buffer and 10 U BseRI (New England Biolabs) were added and incubated at 37 °C for 1 h 50 min. The vector (15 μg) was digested under the same conditions as described above in a total volume of 300 μl and then separated by agarose gel electrophoresis (0.8% w/v agarose, Chapter 2.11) and purified (Chapter 2.6). To avoid re-ligation of the vector, the linearised vector (4 μg) was dephosphorylated in a reaction containing 8 μl 10x arctic phosphatase buffer (New England Biolabs) and 15 U arctic phosphatase (New England Biolabs) in a total volume of 80 μl for 1 h at 37 °C followed by heat inactivation at 65 °C for 5 min. The ligation of the lamATR2tr or licATR2tr insert and the LIC-vector was done in the ratios of 2:1 and 4:1. A total volume of 3 μl contained 0.5 μl 10x T4 DNA ligase reaction buffer and 120 U T4 DNA ligase (New England Biolabs) and the reaction was incubated overnight in a 21 water bath with a temperature starting at 21 °C lowered to 4 °C. The ligation product was transformed into E. coli DH5α (Chapter 2.8.4), transformants screened on selective agar containing kanamycin and verified using the primer Lic_BseRI_F, lam_BseRI_F1 and Atr2_NdeI_Lic_R (Table 5.1). Positive clones were incubated overnight for plasmid preparation (Chapter 2.5) and a restriction digest performed with BseRI and NdeI (Chapter 2.8.6). For final confirmation that the right plasmid with the correct sequence was engineered, the gene was sequenced (Chapter 2.8.9) using T7 and T7term primers.

5.3.2 Cloning of P450s into lamATR2tr- and licATR2tr-vector

5.3.2.1 Primer design

In cloning experiments, P450 genes were cloned using either their native sequence or their codon-optimised sequence for expression in E. coli. Codon optimisation was performed by GeneArt (Table 5.2). The full length sequences are found in Appendix B.
Table 5.2: List of P450s used for the creation of artificial plant fusions with the reductase ATR2tr

<table>
<thead>
<tr>
<th>P450</th>
<th>gene origin</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>IFS</em>&lt;sub&gt;native&lt;/sub&gt;</td>
<td>construct containing the native gene derived from Prof. Mattheos Koffas&lt;sup&gt;306&lt;/sup&gt;</td>
</tr>
<tr>
<td><em>IFS</em>&lt;sub&gt;opt&lt;/sub&gt;</td>
<td>synthesized and codon optimized for <em>E. coli</em> expression by GeneArt</td>
</tr>
<tr>
<td>CYP73A5</td>
<td>synthesized and codon optimized for <em>E. coli</em> expression by GeneArt</td>
</tr>
<tr>
<td>CYP82E4</td>
<td>native gene derived from Prof. Ralph Dewey&lt;sup&gt;427&lt;/sup&gt;</td>
</tr>
<tr>
<td>CYP81D8</td>
<td>native gene isolated from <em>Arabidopsis thaliana</em></td>
</tr>
</tbody>
</table>

To each forward primer the sequence CCAGGGACCAGCA was added upstream of the start codon of the P450 sequence to create the overhangs for the ligation independent cloning. The stop codons of the P450 inserts were removed in order to obtain a fusion to ATR2tr. The reverse primers were designed individually for each linker (Table 5.3). To prepare the sequence for T4 treatment, as described previously, the 3’-end of the sequence was engineered to end with an Adenine (A), resulting in a change of the last amino acid: in CYP73A5 the terminal Cysteine was changed to a Serine and in CYP82E4 the terminal Tyrosine to a Threonine. The last amino acids of IFSori, IFSsyn and CYP81D8 were silent mutations (see Appendix B).
Table 5.3: Primer sequences for cloning P450 inserts into ATR2tr-LIC-vector
Lic primer sequence addition in capital letters, insert sequence in small letters

<table>
<thead>
<tr>
<th>primer name</th>
<th>sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>e-IFS for ATR2 fusion F</td>
<td>CCAGGGACCAGCA atggctctgttattageagttttcttgtgtt</td>
</tr>
<tr>
<td>eIFSori-LicATR2tr R</td>
<td>GAGGAGAAGGCGCGtgaagaggttagatgaecgecgatctgc</td>
</tr>
<tr>
<td>eISFori-lamLicATR2tr R</td>
<td>GAGGAGCAGGAGCCtgaagaggttagatgaecgecgatctgc</td>
</tr>
<tr>
<td>eIFSsyn-native-Lic F</td>
<td>CCAGGGACCAGCA atggcactgctgtgcagttttctgtt</td>
</tr>
<tr>
<td>IFSSyn-trunc-ATR2tr F</td>
<td>CCAGGGACCAGCA gccegtccgaccggcagcagctcactcag</td>
</tr>
<tr>
<td>73A5syn-trunc-Lic F</td>
<td>CCAGGGACCAGCA etcggtggtaaaaactggaatgctcggggct</td>
</tr>
<tr>
<td>73A5syn-LicATR2tr R</td>
<td>GAGGAGAAGGCGCG tgaagctcggtgccacaaatcatctgtatg</td>
</tr>
<tr>
<td>73A5syn-lamLicATR2tr R</td>
<td>GAGGAGCAGGAGCC tgaagctcggtgccacaaatcatctgtatg</td>
</tr>
<tr>
<td>82E4-trunc-Lic F</td>
<td>CCAGGGACCAGCA acacaatactgctcaaanaccacctttaac</td>
</tr>
<tr>
<td>82E4-LicATR2tr R</td>
<td>GAGGAGAAGGCGCG tgaagctcggtgccagcaggecgctattatc</td>
</tr>
<tr>
<td>82E4-lamLicATR2tr R</td>
<td>GAGGAGCAGGAGCC tgaagctcggtgccagcaggecgctattatc</td>
</tr>
<tr>
<td>81D8-trunc-Lic F</td>
<td>CCAGGGACCAGCA gggaaaactcaagegaaaacgataaccctccgag</td>
</tr>
<tr>
<td>81D8-LicATR2tr R</td>
<td>GAGGAGAAGGCGCG tcaagctgggtgaaaggttaaaacccagggac</td>
</tr>
<tr>
<td>81D8-lamLicATR2tr R</td>
<td>GAGGAGCAGGAGCC tcaagctgggtgaaaggttaaaacccagggac</td>
</tr>
</tbody>
</table>

5.3.2.2 Cloning strategy
The P450 inserts were amplified by PCR using the high fidelity Phusion polymerase. Primers are listed in Table 5.3 and 4 M NDSB 201 (Chapter 2.8.2). The purified P450 sequences (Chapter 2.6) as well as lamATR2tr- and licATR2tr-vector were prepared following the protocol for ligation independent cloning (Chapter 1.1). Primer T7 and ATR2seqR (Table 2.4) were used for sequence confirmation (Chapter 2.8.9).

5.3.3 Expression and activity assays of the novel plant fusions
The expression of the plant P450 fusion was performed overnight at 20 °C following the protocol described in Chapter 2.10.1 and activity assay in Chapter 2.15.3.
5.4 Results

5.4.1 Developing the platform technology for plant P450-ATR2tr fusions

5.4.1.1 Designing the linker region between P450 and reductase

Two different linkers between the P450s and the reductase ATR2tr were tested (Figure 5.4).

Figure 5.4: Schematic construction of the P450-ATR2tr fusion containing the lam-linker

The first linker to be tested was the lic-linker (amino acid sequence: RAFSS), which is the necessary overhang in the LIC-vector for the LIC technology described in Chapter 1.1. The second linker was the λ-linker (amino acid sequence: GST) created by the Koffas group. The λ-linker and the first five
amino acids of the truncated CPR (from *C. roseus*) resulted in the sequence of GSTSSGSG (called lam-linker), which was modified by inserting a *Bse*RI restriction site. The codon for the first glycine and for threonine had to be changed to be compatible for the ligation independent cloning. Threonine is coded through four possible codons (ACN, where N is A, T, G or C), which could not be used, because the linker sequence is part of the overhangs created through the T4 polymerase reaction and therefore must not contain the base adenine. Threonine was changed to the similar, neutral, polar amino acid serine and the codon for glycine from GGA to GGC in a silent mutation to remove the adenine from the nucleotide sequence. However, these changes resulted in four subsequent serines and the formation of two *Bse*RI (CTCCTC) sites. Therefore, the threonine was changed to cysteine.

### 5.4.2 Cloning of ATR2tr into the LIC-vector

The ATR2tr with the lam-linker or the lic-linker at the N-terminus was cloned into the LIC-vector. The lamATR2tr- and licATR2tr-vectors were transformed into *E. coli* DH5α and positive colonies screened using whole cell PCR (Chapter 2.8.2). Plasmids of positive clones were purified and digested to verify the correct insert (Figure 5.5).

![Figure 5.5: Analysis of positive transformants for the presence of the lam-ATR2tr in the LIC-vector](image)

Restriction digest of the LIC-vector (negative control) and lamATR2tr by BseRI (*Bse*R), NdeI (*Nde*) or with both enzymes (2cut)
5.4.3 Cloning of P450 inserts into lamATR2tr and licATR2tr

The P450 inserts were amplified by PCR using the high fidelity Phusion polymerase and cloned into the lamATR2tr- and the licATR2tr-vector. Again colonies containing the desired gene were detected by whole cell PCR and correct inserts were confirmed by restriction digest of the (Figure 5.6) and sequencing.

![Figure 5.6: Analysis of positive transformants for the presence of the P450 73A5tr]

Restriction digest with EcoR1 (Eco) of the empty lamATR2tr-vector and the transformants 73A5tr-lamATR2tr No1 and 2 containing the insert and No3 with no insert.

5.4.4 Expression and activity assays of the novel plant fusions

5.4.4.1 IFS-ATR2tr fusions

The isoflavone synthase activity was detected for the conversion of naringenin to genistein by HPLC after extraction of the culture supernatant in ethylacetate (Chapter 2.15.3.3). Great error bars were received due to the complex extraction method and an internal standard was required. Three compounds (syringic acid, biochanin A and scopoletin) were chosen to test for their use as internal standard for the detection of naringenin and genistein by HPLC (Chapter 2.15.3.3). Retention time and absorption maxima for each compound are listed in Table 5.4.
Table 5.4: List of chemical compounds considered to use as internal standard optimising the extraction method for the two flavonoids naringenin and genistein

<table>
<thead>
<tr>
<th>Chemical compound (IUPAC name)</th>
<th>Molecular weight (g/mol)</th>
<th>Absorption maximum (nm)</th>
<th>Retention time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naringenin (5,7-dihydroxy-2-(4-hydroxyphenyl)chroman-4-one)</td>
<td>272</td>
<td>291</td>
<td>12.0</td>
</tr>
<tr>
<td>Genistein (5,7-Dihydroxy-3-(4-hydroxyphenyl)chromen-4-one)</td>
<td>270</td>
<td>261</td>
<td>16.8</td>
</tr>
<tr>
<td>Syringic acid (4-hydroxy-3,5-dimethoxybenzoic acid)</td>
<td>198</td>
<td>276</td>
<td>4.2</td>
</tr>
<tr>
<td>Biochanin A (5,7-dihydroxy-3-(4-methoxyphenyl)chromen-4-one)</td>
<td>284</td>
<td>260</td>
<td>16.8</td>
</tr>
<tr>
<td>Scopoletin (7-Hydroxy-6-methoxycoumarin)</td>
<td>192</td>
<td>230 &amp; 346</td>
<td>4.8</td>
</tr>
</tbody>
</table>

Syringic acid could be not detected after extraction, however it was stable at room temperature in methanol over days. Syringic acid was lost during the extraction, due to the deprotonation of the carboxyl group and so it was not extracted into the organic phase of ethyl acetate. Therefore, syringic acid was not used as internal standard when using this method.

Biochanin A is a homologue of genistein and therefore, expected to have similar properties. Unfortunately, biochanin A is not very stable at room temperature and has the same retention time of 16.8 min as genistein (Figure 5.7) and therefore unsuitable as internal standard.

Figure 5.7: HPLC analysis of the flavonoids biochanine A, genistein and naringenin
HPLC chromatogram at 260 nm for a sample mix of biochanin A (50 μM, retention time: 16.8 min), naringenin (50 μM, retention time: 12.0 min) and genistein (50 μM, retention time: 16.8 min); HPLC conditions: 50 μl sample was applied to HPLC, TechSphere column, 1 ml/min isocratic flow 50% water, 50% methanol with 0.1% acetic acid, column temperature: 25 °C, sample temperature: 21 °C, run time 23 min
Scopoletin elutes earlier than naringenin and genistein with a retention time of 4.8 min (Figure 5.8). It is stable at room temperature over days and also during the extraction step and was therefore chosen as internal standard for the isoflavone synthase activity assay (Chapter 2.15.3.3).

Figure 5.8: HPLC analysis of the flavonoids scopoletin, naringenin and genistein
HPLC chromatogram at 280 nm for scopoletin (50 μM, retention time: 4.8 min), naringenin (50 μM, retention time: 12.0 min) and genistein (50 μM, retention time: 16.8 min); HPLC conditions: 50 μl sample was applied to HPLC, TechSphere column, 1 ml/min isocratic flow 50% water, 50% methanol with 0.1% acetic acid, column temperature: 25 °C, sample temperature: 21 °C, run time 23 min

5.4.4.1.1 Activity in a growing cell assay

The isoflavone synthase I activity of the different artificial fusions was tested in a growing cell assay. The conversion of the substrate naringenin (50 μM) to genistein was observed by analysing samples of the culture supernatant over time. No genistein was produced by the recombinant expressed fusions containing the native \( IFS \) gene (Figure 5.9). For the codon optimised \( \text{IFSopt-lamATR2tr} \) and \( \text{IFSopt-licATR2tr} \) 5 μM (10% conversion) and 10 μM (20% conversion) genistein were detected after 48 h, respectively. No activity was found for the negative controls \( IFS \) (in the LIC-vector, data not shown), empty lamATR2tr or empty licATR2tr.
Figure 5.9: Activity of *Glycine max* IFS fused to Arabidopsis ATR2tr

Genistein production in a growing cell assay by artificial plant P450 fusions (IFSnat = native isoflavone synthase I from *G. max*, IFSopt = isoflavone synthase I codon optimized for *E. coli* expression, lamATR2tr = truncated ATR2 with lambda linker, licATR2tr = truncated ATR2 with lic-linker). The error bars represent the mean of five independent replicas ± standard deviation.
The fusions containing the IFS and ATR2tr were analysed by SDS PAGE and western blot analysis using anti His-antibodies (Chapter 2.14.1) and anti ATR2-antibodies (Chapter 2.14.2). No signal was received for the fusions using anti His-antibodies, possibly a result of a not accessible His-tag.

The IFS fusions containing the native gene did not show a signal, possibly because expression levels were below detection limits.

The IFSopt-lamATR2tr and IFSopt-licATR2tr were detected only in the insoluble fraction (Figure 5.10) and were identified by MALDI-MS analysis (Chapter 2.15.2). In all sample (total protein, soluble and insoluble protein fraction) was found ATR2tr (72 kDa) implying a degradation of the fusions. For positive control was used lamATR2tr without IFS.

**Figure 5.10:** A) SDS-PAGE and B) western blot analysis (anti ATR2-antibodies) of *Glycine max* IFS fused to Arabidopsis ATR2tr

Analysis of total protein, soluble and insoluble protein (1 = IFSopt-lamATR2tr, 2 = IFSopt-licATR2tr, 3 = lamATR2tr control). The protein band with the expected size of IFSopt-lamATR2tr and IFSopt-licATR2tr (red boxes) from the insoluble protein fraction were sent for Maldi-MS analysis.
5.4.4.2 CYP73A5tr-ATR2tr fusions

5.4.4.2.1 Activity in a resting cell assay

The CYP73A5 codon optimised gene sequence without the hydrophobic N-terminus (73A5tr) was fused to ATR2tr using two different linkers. Both fusions, 73A5tr-lamATR2tr and 73A5tr-licATR2tr, converted cinnamic acid to coumaric acid in a resting cell assay, when grown in LB medium. The activity was three to five times greater when the fusions were grown in M9 minimal medium than in LB medium (Figure 5.11). No product was detected for the empty-ATR2tr controls, independent of the used linker.

![Graph showing activity of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr in different media.](image)

**Figure 5.11: Activity of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr in different media**

Production of coumaric acid by 73A5tr-lamATR2tr and 73A5tr-licATR2tr in a resting cell assay after expression in M9 minimal or LB medium at 15 °C and 200 rpm.

Different expression conditions were tested in M9 medium to increase the activity of the CYP73A5-fusions. When fusions were expressed shaking (200 rpm) at 15 °C, resting cell assay showed more than 30% of the cinnamic acid was converted to coumaric acid in 24 h. Preliminary results indicated that the 73A5tr-licATR2tr produced approximately 5% more coumaric acid suggesting slightly higher activity than 73A5tr-lamATR2tr.
Activity was significantly reduced when cells were grown in M9 medium shaking (200 rpm) at 20 °C and 200 rpm (Figure 5.12) in comparison to the result above (Figure 5.11).

Figure 5.12: Activity of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr in M9 medium
Production of coumaric acid by 73A5tr-lamATR2tr and 73A5tr-licATR2tr in a resting cell assay after expression in M9 minimal medium at 20 °C and 200 rpm. The error bars represent the mean of five independent replicas ± standard deviation.
Similar amounts of coumaric acid product (30%) were detected for 73A5tr-lamATR2tr when expressed at 15 °C and 50 rpm (Figure 5.13) compared to 15 °C and 200 rpm. The fusion 73A5tr-licATR2tr only converted substrate to 20% coumaric acid at 15 °C and 50 rpm, which is 50% of the conversion received at 15 °C and 200 rpm (Figure 5.11).

Figure 5.13: Conversion of cinnamic acid to coumaric acid by artificial plant fusions containing Arabidopsis P450 73A5tr and Arabidopsis ATR2tr in a resting cell assay
Protein expression at 15 °C and 50 rpm (cinn = cinnamic acid, coum = coumaric acid, 73A5tr = truncated version of CYP73A5, synthesised and codon optimised for the expression in E. coli, lamATR2tr = truncated version of ATR2 with lam-linker, licATR2tr = truncated version of ATR2 with lic-linker, control-lamATR2tr and control-licATR2tr = empty vector control containing the ATR2tr with different linkers). The error bars represent the mean of four independent replicas ± standard deviation.

5.4.4.2 Purification of 73A5tr-lamATR2tr
The fusion 73A5tr-lamATRtr (127 kDa) was subjected to nickel-affinity chromatography (Chapter 2.10.5) and then detected by SDS PAGE and western blot analysis (Figure 5.14). The 73A5tr-lamAR2tr was mainly present in the insoluble and in the eluted protein fraction. The protein band of the expected size for 73A5tr-lamATR2tr was sent for MALDI-MS analysis (Chapter 2.15.2) and the expected two proteins cinnamte-4-hydroxylase (from Arabidopsis) and NADPH-ferrihemoprotein reductase (Arabidopsis) were identified.
The different fractions were then tested for the ability to convert cinnamic acid. Results showed that activity was only found in the insoluble fraction (Figure 5.15).

To confirm this result, 73A5tr-lamATR2tr was expressed again in *E. coli* Rosetta 2 (DE3) and fractions of whole cells, sonicated cells (total protein), supernatant (soluble protein) and dissolved pellet (insoluble fraction) after sonication and centrifugation were tested for activity and analysed by SDS PAGE and western blot analysis. The activity found in whole cells (cell concentration:
50 mg/100ml, Figure 5.16A) with 170 μM coumaric acid was reduced to 10% in total and to 3% in the insoluble protein fraction (Figure 5.16B).

**Figure 5.16:** Cinnamate-4-hydroxylase activity of 73A3tr-lamATR2tr in different fractions after expression in *E. coli*

Production of coumaric acid by 73A5tr-lamATR2tr in A resting cells and B different fractions after the sonication

The 73A5tr-lamATR2tr fusion was detected by western blot analysis in the whole cells, total and insoluble protein fraction (Figure 5.17). No signal was seen in the soluble fraction.

**Figure 5.17:** SDS-PAGE and western blot analysis of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr as soluble and insoluble protein

73A5tr-lamATR2tr fusion (127 kDa) and lamATR2tr (negative control) detected by A SDS PAGE and B western blot analysis against anti His-antibodies (1 = whole cells, 2 = total protein fraction, 3 = fraction of soluble protein, 4 = fraction of insoluble protein)
5.4.4.3 CYP82E4tr-ATR2tr fusions

The N-demethylase (CYP82E4) was truncated to remove the membrane anchor (sequence, see Appendix B) and cloned into the lamATR2tr- and the licATR2tr-vector. The expression was performed in *E. coli* Rosetta2 (DE3) and M9 minimal medium (Chapter 2.10.1). The activity of the fusion was tested using resting cell assays following the conversion of nicotine to nornicotine (Chapter 2.15.3.5). Nicotine was removed from all samples, including negative controls (control-lamATR2tr, Rosetta 2 cells, and no cell control) within 45 min (Figure 5.18). The disappearance of nicotine was found to be due to the instability of nicotine in 70% water and 30% methanol (0.2% phosphoric acid, pH 7.25 with triethylamine).

![Figure 5.18: Nicotine removal in a resting cell assay](image)

The error bars represent the mean of three independent replicas ± standard deviation.

Nevertheless, the reaction mixture was analysed to see if nornicotine was produced. Nornicotine standard had a retention time of 4.2 min. Nornicotine could not be detected in any of the samples, due to other products of *E. coli* eluting between the retention time of 3.3 to 5.7 min and therefore was not detected (Figure 5.19).
Figure 5.19: HPLC analysis (reversed-phase column) of the crude cell extract for nicotine and nornicotine.
HPLC chromatogram for A 82E4tr-lamATR2tr and B control-lamATR2tr after a reaction time of 10 min in a resting cell assay (17.2 min = nicotine), HPLC conditions: SunFire C18 column, 1 ml/min flow isocratic 70% water (0.2% phosphoric acid, pH 7.25 with triethylamine) and 30% methanol (0.2% phosphoric acid pH 7.25 with triethylamine), column temperature: 20 °C, sample temperature: 4 °C.

To improve the identification of nornicotine, a chiral column (ChiralPak IA) was used. This column was able to separate a commercial racemic mixture of nornicotine (Figure 5.20). The enantiomers of nicotine eluted in one peak at 15.3 min. However, when extracts were tested, it was not possible to isolate nicotine or nornicotine using this method.
Figure 5.20: HPLC analysis (chiral column) of the crude cell extract for nicotine and nornicotine

HPLC separation of nicotine (retention time: 15.3 min), (R)-nornicotine (9.9 min) and (S)-nornicotine (11.7 min), HPLC conditions: ChiralPak IA column, isocratic 0.5 ml/min flow of 35% ethanol and 65% (heptanes : isopropanol : TFA = 90:10:0.1), 1 mM (R/S)-nicotine, 1mM (R/S)-nornicotine

Both fusions, 82E4tr-lamATR2tr and 82E4tr-licATR2tr (128 kDa, red boxes in line 3 and 4, respectively, Figure 5.21), were detected by SDS PAGE and western blot analysis (with anti His-antibodies) in the insoluble fraction.

Figure 5.21: SDS-PAGE and western blot analysis of Nicotiana tabacum P450 82E4tr fused to Arabidopsis ATR2tr

Analysis of crude extract by SDS-PAGE A in soluble fraction and B in insoluble fraction as well as by western blot analysis against anti His-antibodies C in soluble fraction and D in insoluble fraction (1 = control-licATR2tr, 2 = Rosetta 2 cells, red boxes: 3 = 82E4tr-lamATR2tr, 4 = 82E4tr-licATR2tr)
5.4.4.4 CYP81D8tr-ATR2tr fusions

The 81D8tr-lamATR2tr and 81D8tr-licATR2tr were expressed in Rosetta 2 (DE3) and tested for activity against 100 μM TNT. There were no differences in TNT transformation between the fusion proteins (81D8tr-lamATR2tr and 81D8tr-licATR2tr) and the negative control (Figure 5.22) and no additional peaks were seen.

![Figure 5.22: TNT removal in a resting cell assay over time](image)

The error bars represent the mean of four independent replicas ± standard deviation.

As the endogenous function of CYP81D8 is unknown, the 81D8tr-fusions were also tested with 7-ethoxycoumarin, a model P450 substrate, in resting cell assays. The concentration of 7-ethoxycoumarin remained constant (Figure 5.23) and no product peak appeared over time.
The 81D8tr-lamATR2tr (125 kDa) and 81D8tr-licATR2tr (125 kDa) were detected in the insoluble fraction by SDS PAGE and western blot analysis (Figure 5.24). No signal was visible in the soluble fraction except for control-licATR2tr (72 kDa).

Figure 5.23: Resting cell assay using 7-ethoxycoumarin as substrate
The error bars represent the mean of four independent replicas ± standard deviation.

Figure 5.24: SDS-PAGE and western blot analysis of Arabidopsis P450 81D8tr fused to Arabidopsis ATR2tr
Detection of proteins by SDS A in soluble fraction and B in insoluble fraction as well as by western blot analysis against anti His-antibodies C in soluble fraction and D in insoluble fraction (1 = control-licATR2tr, 2 = Rosetta 2 cells, red boxes: 3 = 81D8tr-lamATR2tr, 4 = 81D8tr-licATR2tr)
5.5 Discussion

A platform for expression of plant P450s fused to plant P450 reductases was developed by cloning the Arabidopsis reductase ATR2tr. ATR2tr was expressed in a soluble form in *E. coli* and chosen for the development of the platform technology due to its higher activity than other plant reductases, such as Arabidopsis ATR1tr, ferredoxin reductase and cytochrome c reductase (Chapter 4).

ATR2tr was cloned into the LIC-vector using traditional cloning methods. The connection between the P450 and the reductases was realised by a linker. Two different linkers (lic-linker and lam-linker) were tested. To test this platform, four plant P450s catalysing different reaction types were selected as candidate P450s. The hydrophobic membrane anchor regions of the selected P450s, as identified using TMHMM and SignalP3.0 software (Chapter 2.8.1) were removed and P450s (except CYP93C1 was used with a modified N-terminus) cloned into lamATR2tr- and licATR2tr-vector using ligation independent cloning, expressed and assayed.

Expression of the artificial plant fusions in the *E. coli* host did not affect the cell growth under the conditions tested here, although it was reported that the fusion of CYP73A4 (cinnamate-4-hydroxylase from *C. roseus*) to the reductases from *C. roseus* resulted in reduced *E. coli* growth.\(^{299}\)

IFS-ATR2tr fusions

The isoflavin synthase (IFS) was first identified after functional expression in yeast and Arabidopsis.\(^{428}\)

The activity of IFS from *Glycine max* fused to ATR2tr was tested in a growing cell assay for the conversion of naringenin to genistein. No genistein was produced using the native IFS gene (IFSnat-lamATR2tr and IFSnat-licATR2tr), however a conversion of 10% and 20% was seen after 48 h for the codon optimised IFSopt-lamATR2tr and IFSopt-licATR2tr, respectively. The fusion IFSsyn-lamATR2tr converted 5.9 μM cinnamic acid to coumaric acid in resting cell assays after 18 h. Other research groups produced similar amounts of genistein (3.7 μM) with yeast microsomes containing IFS after 10 h.\(^{428}\)

An example of another artificial plant fusion of rcIFS-riceCPR (IFS from red clover fused to rice reductases) expressed in *E. coli* produced 63% genistein after
12 h in a growing cell assay (using LB medium), which was three times higher than when IFS fused to ATR2tr (this project).

**73A5tr-ATR2tr fusions**

The Arabidopsis CYP73A5 hydroxylates cinnamic acid to coumaric acid. Both fusions (73A5tr-lamATR2tr and 73A5tr-licATR2tr) were expressed in *E. coli* and activity was shown in resting cell assays. The conversion to coumaric acid was increased by altering the growth conditions prior to the resting cell assay. Expression overnight in M9 medium resulted in five fold higher activity (after 24 h cell assay). The opposite effect was demonstrated for rcIFS-riceCPR expressed in *E. coli*, where the production of genistein was reduced to 35% when M9 instead of LB medium was used. At a yield of 35% genistein after 48 h the substrate naringenin is not used up (still 50% remained), which could be a result of low enzyme activity when the substrate concentration is too low. This can be tested by adding again naringenin. Another reason of this phenomenon could be the consumption of the cofactor NADPH, which is only replaced when a carbon source in *E. coli* is available. Resting cell assays were usually performed in 50 mM potassium phosphate buffer (Chapter 2.15.3.4) without any additional carbon source. If the NADPH amount is the limiting factor can be tested by adding sugar, which will be transferred into the cell and transformed in the pentose phosphate pathway to create NADPH.

Furthermore, in this project it was found that an expression at 15 °C and 200 rpm increased the conversion of more than 30%, whereas the activity of 73A5tr-licATR2tr was 5% higher than for 73A5tr-lamATR2tr.

The fusion 73A5tr-lamATR2tr was subjected as soluble protein to nickel-affinity chromatography to prove the presence of the fusion protein as confirmed in western blot and MALDI-MS analysis, however no activity was detected. Activity studies and western blot analysis of the different fractions (whole cells, total, soluble and insoluble fraction) indicated that 73A5tr-lamATR2tr was present as membrane associated protein and therefore found in the insoluble fraction. CYP73A4 from *C. roseus* (a homolog of CYP73A5) and CYP71D12 (tabersonine 16-hydroxylase from *C. roseus*), both fused to the reductases from *C. roseus* were also found as active proteins in the membrane fraction. These observations
indicate that the association of the P450 with the endoplasmic reticulum is necessary for activity.

82E4tr-ATR2tr fusions

CYP82E4 from Nicotiana tabacum catalyses the N-demethylation of nicotine to nornicotine and both could be separated by HPLC. The novel fusions (82E4tr-lamATR2tr and 82E4tr-licATR2tr) were expressed in E. coli and activity tested in resting cell assays. The substrate nicotine was removed for all tested samples (also the negative controls) in less than 1 h, which indicated that nicotine is not very stable under the reaction conditions. Moreover, a product peak for nornicotine could not be definitive allocated due to the presence of interfering peaks from E. coli extract, which appeared around the same retention time as nornicotine. Other research groups have avoided this problem by using $^{14}$C-labelled nicotine as substrate for thin layer chromatographic analysis, which is a more sensitive method.\(^{427,429-431}\)

Nicotine and its derivates were shown to be separated from urine using HPLC (reversed-phase C18 column).\(^{432,433}\) However, when a chiral column was used to identify nornicotine by adapting the method of Demetriou et al. 1993\(^{434}\) no improved separation could be seen when resting cell assay samples were separated. Further attempts to optimise this method were not made due to time constraints.

These preliminary were unable to confirm whether fusions containing the CYP82E4 is active. In the future, GC-MS analysis could be used to detect both nicotine and nornicotine.

81D8tr-ATR2tr fusions

The Arabidopsis CYP81D8 was 30 times upregulated in a microarray experiment in Arabidopsis plants after TNT treatment.\(^{363}\) While activity towards TNT by the fusions (81D8tr-lamATR2tr and 81D8tr-licATR2tr) was not detected, when 7-ethoxycoumarin was tested as potential substrate also no product appeared. So whether the fusions were active has not yet been confirmed. E. coli can use TNT as nitrogen source.\(^{435-437}\) and TNT was reduced to hydroxylaminodinitrotoluenes in all samples by the nitroreductases naturally present in E. coli.
Two linkers (lic-linker with the amino acid sequence: RAFSS and lam-linker: GSTSSGSG) have been tested to create fusions of the P450s and the reductase ATR2tr. Generally, the lam-linker resulted in 5-10% higher activity than the lic-linker when tested with IFS or 73A5 fusions. However, preliminary results suggested that the 73A5tr-licATR2tr when grown in M9 medium at 15 °C and 200 rpm may have similar activity to 73A5tr-lamATR2tr.

The artificial lam-linker was originally engineered by Prof. Koffas’ research group to create a junction of GSTSSGSG with the start of the truncated CPR, which should avoid the formation of secondary structure. In their work, they compared the activity of IFS-CPR connected with the lam-linker and the construct C, a fusion without a linker, where the IFS C-terminus was directly fused to the truncated CPR from *C. roseus*. Construct C did not produce any product when expressed in *E. coli* JM109, TOP10F’ or BL21Star. However, construct C showed activity when expressed in DH5α, which was 50% compared to IFS-CPR. Their findings showed the necessity of a linker to keep a distance between P450 and reductase. The same GST-linker was used successfully for the artificial plant fusion rcIFS-riceCPR (IFS from red clover fused to rice reductases) as rcIFS-CPR (IFS to *C. roseus* reductases) to allow the two enzymes a separate protein folding.

Previously, Prof. Schröder’s research group also used a ST-linker for engineering different plant fusions with the CPR from *C. roseus*. The dipeptide ST was chosen to avoid forming of extensive secondary structures, which can affect the electron transfer between the reductase and the P450.

The detection of artificial plant fusion enzymes by western blot analysis seemed to be not straightforward, due to lack of data in the literature. An active fusion of CYP73A4-CPR was only detected when containing the membrane anchor of the P450, but then it was not active. Activity was found when the first 69 amino acids of the P450 have been removed, however no protein signal was visible in the immunoblot. The flavonoid hydroxylase (CYP75 from *C. roseus*) was detected as fusion with the reductase from *C. roseus* and the reductase from *Petunia hybrida* by western blot analysis.

All fusions engineered with the Arabidopsis ATR2tr reductase were detected in the western blot analysis using anti His-antibodies. They were found mainly in the
insoluble protein fraction, which implied that these fusions are still membrane associated, despite lacking the hydrophobic N-terminus. The only exceptions with no signal in the SDS PAGE and the western blot were the IFS fusions containing the native IFS gene (IFSnat-lamATR2tr and IFSnat-licATR2tr), for which also no activity was detected, possibly caused by a too low protein expression. The presence of plant fusion proteins containing IFS from G. max or IFS from red clover was not confirmed so far.\textsuperscript{306,307}

This development of a vector, which contained the soluble Arabidopsis reductase ATR2tr, is usable for high-throughput cloning due to the Ligation Independent Cloning system. ATR2tr, one of the Arabidopsis reductase, would be expected to be an appropriate partner enzyme for crating fusion with the 178 Arabidopsis P450s with unknown function to discover their native role in the plant. This system provides a platform technology, where also other plant P450s can be cloned fast and easily as fusion enzymes, which then can be studied further for their potential function.
6 Comparison of plant Cytochromes P450 fused to different reductases

6.1 Introduction

The electron transfer from P450 reductases to P450s is species unspecific, in all cases examined; for example *Nicotiana tabacum* was transformed with the rabbit liver P450 and showed to be supported endogeneous *N. tabacum* reductases. However, the specific activity of the membrane associated CYP75 (flavonoid hydroxylase) from *Petunia hybridia* fused to its own reductase showed twice much activity after recombinant expression in *E. coli* than when fused to the reductase of *C. roseus*. Thus, it could be predicted that the reductases in each organism should be the most efficient for the support of their corresponding P450s considering that they have evolved together.

The CPR from *C. roseus*, firstly characterised from Madyastha and Coscia in 1979 after purification from five day old, etiolated seedlings, was used as the partnering reductase for the creation of artificial plant P450 fusion systems (see Table 1.2, page 45). For all the plant-plant fusions, the hydrophobic N-terminus of the CPR was deleted to avoid an association with the bacterial membrane and to increase the solubility.

The truncated form of ATR2 (ATR2tr) from Arabidopsis was expressed as a soluble and active reductase in *E. coli* (Chapter 4, page 111) and then used as partner enzyme to develop a platform technology of a plant P450 fused to a Arabidopsis reductase (Chapter 5, page 137).

The RhF reductase is originally part of the native fusion P450 RhF from *Rhodococcus* sp. which catalyses the O-dealkylation of 7-ethoxycoumarin to 7-hydroxycoumarin (Figure 1.12). Recently, it was used for the functional expression of bacterial P450 fusions.

6.2 Objectives

A technology platform comprising plant P450s fused to the truncated Arabidopsis reductase ATR2tr was engineered as described in Chapter 5 and in this chapter, the system was developed further by testing additional reductases and *E. coli* strains. The reductases were fused to: cinnamate-4-hydroxylase (CYP73A5) from...
Arabidopsis and isoflavone synthase I (IFS) from *Glycine max* and the activities *in vivo* compared.

The IFS (CYP93C) gene was fused to three reductases:
1) the CPR from *C. roseus* (fusion construct obtained from Prof. Mattheos Koffas, Chapter 5.1),
2) the ATR2tr from Arabidopsis (developed in Chapter 5) and
3) the RhF reductase from *Rhodococcus* sp.

The IFS activity was tested by measuring the conversion of naringenin to genistein in growing and resting cell assays.

The cinnamate-4-hydroxylase (CYP73A5) gene was fused to two reductases:
1) the ATR2tr from Arabidopsis (developed in Chapter 5) and
2) the RhF reductase from *Rhodococcus* sp.

CYP73A5tr from Arabidopsis was chosen as its function in the phenylpropanoid pathway is known. The activity of CYP73A5tr was tested for the hydroxylation of cinnamic acid to coumaric acid in *in vitro* as well as in resting cell assays.

### 6.3 Methods

#### 6.3.1 Expression of IFS-fusions

##### 6.3.1.1 Expression conditions of IFS-CPR

The construct IFS-CPR (C-terminus of isoflavone synthase I from *G. max* fused to the N-terminus of the reductases from *C. roseus*) was kindly provided by Professor Koffas (University at Buffalo, The State University of New York) and was used as positive control for the IFS-P450-reductase fusion systems. Two different *E. coli* strains were tested: JM109, used for expression of IFS-CPR by Leonard and Koffas 2007, and Rosetta 2 (DE3), which contains an additional plasmid for eukaryotic rare codons (chloramphenicol resistance, Novagen) (*E. coli* strains, see Table 2.1). Both *E. coli* strains were transformed with the IFS-CPR fusion construct in the pTrcHis2/LacZ vector (ampicillin resistant, Invitrogen) (Chapter 2.8.4).
For the expression of IFS-CPR, the published protocol from Leonard and Koffas, 2007 was used. Briefly, a preculture of 10 ml LB medium containing 1% glucose and 100 µg/ml ampicillin (additionally 34 µg/ml chloramphenicol for Rosetta 2 cells) was inoculated with a single colony and incubated overnight shaking (200 rpm) at 37 °C. A volume of 100 ml LB containing 100 µg/ml ampicillin (and 34 µg/ml chloramphenicol for Rosetta 2) was inoculated with the preculture to an OD$_{600}$= 0.1 and incubated shaking (200 rpm) at 37 °C. When OD$_{600}$ had reached 0.8, 1 mM IPTG was added and the culture incubated at 30 °C under shaking for 3 h. Cells were then harvested and resuspended to an OD$_{600}$= 0.6 in M9 minimal medium (final volume: 100 ml) containing 6 nM thiamine and 1 mM IPTG (Figure 6.1).

![Flow scheme of the protocol for the expression of IFS-CPR fusion construct according to Leonard and Koffas](image)

**Figure 6.1**: Flow scheme of the protocol for the expression of IFS-CPR fusion construct according to Leonard and Koffas

### 6.3.1.2 Optimised conditions for the expression of the IFS-RhF fusion

The fusion protein IFS-RhF (isoflavone synthase 1 from *G. max* fused to the reductases RhF from *Rhodococcus* sp.) was expressed under the conditions described in Materials and Methods (Chapter 2.10.1), with the following alteration: After inoculation, the main culture was incubated shaking at 28 °C for approximately 16 h. Cultures were then further incubated shaking at 37 °C until the OD$_{600}$ reached 0.6 to 0.8.
6.4 Results

6.4.1 IFS fusions

6.4.1.1 Expression of IFS-CPR
The fusion IFS-CPR (isoflavone synthase I from *G. max* fused to the reductases from *C. roseus*) was expressed in *E. coli* JM109 and Rosetta 2 (DE3) cells. The presence of the plasmid in the strains was confirmed by PCR.

Following the addition of IPTG, the cell density increased for all samples during the following 3 h incubation (Figure 6.2). However, following the addition of naringenin, during the subsequent growing cell assay, IFS-CPR expression in JM109 cells grew two-fold more slowly than the other cultures (Figure 6.2).

![Graph showing cell growth](image)

**Figure 6.2:** Observation of the cell growth of *E. coli* JM109 and Rosetta 2 during the expression of IFS-CPR
Data for IFS-CPR in JM109 and Rosetta 2 (DE3) cells represent the mean of three independent replicas ± standard deviation.

In the cultures of Rosetta 2 (DE3) cells expressing IFS-CPR, 75% of the substrate naringenin was removed from the medium, whereas in the cultures of JM109 cells expressing IFS-CPR was no significant uptake of naringenin (Figure 6.3A), also small amount of 5 µM genistein was produced. The naringenin concentration
increased in the cultures of the untransformed JM109 and Rosetta 2 (DE3) cells, possibly due to water evaporation over time. The product genistein was produced with IFS-CPR in both *E. coli* strains, with the conversion five times higher in Rosetta 2 (DE3) than in JM109 (Figure 6.3B). Rosetta 2 (DE3) was therefore chosen for further experiments.

![Figure 6.3: Activity of *Glycine max* IFS fused to *Catharanthus roseus* CPR when expressed with *E. coli* JM109 and Rosetta 2 cells.](image)

A naringenin and B genistein concentration in the culture supernatant after an expression time of 1 and 24 h (IFS = isoflavone synthase 1 from *G. max*, CPR = P450 reductase from *C. roseus*). The error bars represent the mean of two to three independent replicas ± standard deviation.
6.4.1.2 Expression of IFS-RhF

The expression of IFS-RhF (plant-bacterial fusion) was performed under the same conditions as stated for IFS-CPR (Chapter 6.3.1.1). However, activity of the isoflavone synthase was lower than IFS-CPR: just 2 μM genistein was produced after 24 h (Figure 6.4), which is 12.5 times less than for IFS-CPR.

Figure 6.4: Activity of Glycine max IFS fused to Rhodococcus sp. RhF reductase
*In vivo* conversion of naringenin to genistein by IFS-RhF. The error bars represent the mean of four independent replicas ± standard deviation.
By using the optimised protocol (Chapter 6.3.1.2), the *in vivo* genistein production in growing cells containing IFS-RhF over 24 h was improved 10-fold. Thus, the plant-bacterial fusion IFS-RhF (red in Figure 6.5) produced more genistein (23 μM after 48 h) than the plant-plant fusion IFS-CPR (green in Figure 6.5, 18 μM after 48 h) in the growing cell assay at 20 °C. No activity was detected for the expression of IFS without reductases, the RhF reductases alone, IFS-RhF boiled control or *E. coli* Rosetta 2 (DE3) cells without any plasmid.

![Graph](image)

**Figure 6.5**: Conversion of naringenin to genistein by IFS-RhF and IFS-CPR in a growing cell assay at 20 °C
IFS = isoflavone synthase 1 from *G. max*, CPR = P450 reductase from *C. roseus*, RhF = P450 reductase domain from *Rhodococcus* sp. The error bars represent the mean of five independent replicas ± standard deviation.
When cells were grown at 15 °C, activity of IFS-RhF and IFS-CPR was reduced to around 15 μM genistein for both after 48 h (Figure 6.6). This was 3.2-fold and 2.5-fold lower for IFS-RhF and IFS-CPR, respectively, than when cells were grown at 20 °C.

The activity seen for IFS-RhF was 10-fold higher than IFS-CPR in growing cell assays. Faster growth and an increase in cell mass may explain the rise in genistein production for IFS-RhF. Therefore, the OD₆₀₀ of E. coli Rosetta 2 (DE3) was measured over a time course (Figure 6.7), however no significant difference was observed. The differences in genistein production might be due to the extraction method and the sensitivity in HPLC detection.
Figure 6.7: Observation of the optical density during the expression of IFS-CPR and IFS-RhF at 20 °C in M9 medium
The error bars represent the mean of five independent replicas ± standard deviation.

The level of the conversion of naringenin to genistein was similar for IFS-CPR and IFS-RhF when tested in a resting cell assay (Figure 6.8).

Figure 6.8: Conversion of naringenin to genistein in a resting cell assay
(nar = naringenin, gen = genistein, IFS = isoflavone synthase 1 from G. max, CPR = P450 reductase from C. roseus, RhF = P450 reductase domain from Rhodococcus sp., control = plasmid containing RhF reductase expressed in E. coli.) The error bars represent the mean of five independent replicas ± standard deviation.
6.4.1.3 Comparison of the IFS fusions in their growing cell cultures

The production of genistein by IFS fused to different reductase was compared in growing cell assays (Rosetta 2 (DE3), 20 °C and 200 rpm). The IFS-RhF produced almost 50% genistein after 48 h and was found to have the highest activity in comparison to the other fusions under these conditions (Figure 6.9). IFS-CPR showed the second highest activity with 35% genistein production. Interestingly, activity for the IFS fusions containing the Arabidopsis ATR2tr as reductase partner was only achieved after codon optimisation of the IFS gene. No such modification was required for the other two fusions (IFS-RhF and IFS-CPR). Rosetta 2 (DE3) harbouring the empty vector did not produce coumaric acid and the substrate concentration of cinnamic acid remained unchanged, verifying the activity of the created fusion enzymes.

![Figure 6.9: Production of genistein by IFS fusions in a growing cell assay](image)

The error bars represent the mean of at least four independent replicas ± standard deviation.

(IFS = isoflavone synthase 1 from *G. max*, opt = codon optimised gene for *E. coli* expression, nat = native gene of IFS, ATR2tr = truncated version of P450 reductase from Arabidopsis ATR2, CPR = P450 reductase from *C. roseus*, RhF = P450 reductase domain from *Rhodococcus* sp.)
6.4.1.4 Detection of IFS fusion enzymes

To show the presence of IFS-CPR and IFS-RhF fusions, SDS PAGE and western blot analysis (anti-His antibody) were conducted. Faint signals on western blot were detected for IFS-RhF in the soluble and insoluble fraction (Figure 6.10). No signal was detected for IFS-CPR by western blot analysis, a result of possibly through a very low expression of IFS-CPR by the host cells.

![Figure 6.10: SDS-PAGE and western blot analysis of Glycine max IFS fused to Catharanthus roseus CPR and to Rhodococcus sp. RhF reductase in the soluble and insoluble fraction](image)

Soluble and the insoluble fraction after the sonication: A SDS-PAGE and B western blot (anti-His antibody); (1 = IFS-CPR, 80 kDa; 2 = IFS-RhF, red boxes. 95 kDa, 3 = RhF-vector control, 36 kDa)

The western blot was repeated with specific antibodies raised against the reductase RhF in order to confirm the results obtained using His-antibodies. The IFS-RhF fusion was detected in the soluble protein fraction as well as in the insoluble protein fraction (Figure 6.11). However, no signal difference was detected in the total protein fraction. No degradation product for the fusion (RhF alone, 35 kDa) was seen. The control RhF alone was detected in the total, the soluble and insoluble protein fraction.
Figure 6.11: SDS-PAGE and western blot analysis of *Glycine max* IFS fused to the RhF reductase from *Rhodococcus* sp. in total, soluble and insoluble protein fraction
Total, soluble and the insoluble protein fraction after the sonication: A SDS-PAGE and B western blot (specific RhF antibody); (1 = IFS-RhF, 95 kDa; 2 = RhF control, 35 kDa)

The fusions IFS-lamATR2tr and IFS-licATR2tr could be not detected by western blot analysis using anti His-antibodies. However, both were detected as fusion proteins by western blot analysis against specific ATR2 antibodies in the insoluble protein fraction (Figure 6.12). No signal was found in the soluble protein fraction and also not in the total protein fraction, possibly due to too low concentration. The control ATR2tr was visible in the total, soluble and insoluble protein fraction (line 3 in Figure 6.12B). ATR2tr alone was also detected in samples containing the IFS-fusions suggesting the degradation of the fusion into IFS and ATR2tr.

Figure 6.12: SDS-PAGE and western blot analysis of *Glycine max* IFS fused to Arabidopsis reductase ATR2tr in total, soluble and insoluble protein fraction
Total, soluble and the insoluble protein fraction after the sonication: A SDS-PAGE and B western blot (specific ATR2tr antibody); (1 = IFS-lamATR2tr, 131 kDa; 2 = IFS-licATR2tr, 131 kDa; 3 = lamATR2tr control, 72 kDa)
Reverse transcription PCR (Chapter 2.7) using primers to the IFS gene was used to test if the genes encoding IFS-RhF were transcribed (Figure 6.13). A band at the predicted size for IFS of 1500 bp was seen using cDNA from IFS-RhF expressing cells as *E. coli* Rosetta 2 (DE3) template. No signal was detected for the boiled IFS-RhF control (line b, Figure 6.13).

![Figure 6.13: Reverse transcription PCR for IFS-RhF](image)

*b* = boiled IFS-RhF control, *c* = RhF control, *I* = IFS-RhF
6.4.1.5 Purification of IFS-RhF

Fractions of Ni-chromatography purified Rosetta 2 (DE3) expressed IFS-RhF were analysed by SDS-PAGE and western blot. IFS-RhF fusion was detected in the insoluble (line 2, Figure 6.14) and in the eluted (line 5) fractions.

![SDS-PAGE and western blot analysis of Glycine max IFS fused to Rhodococcus sp. RhF reductase after purification with Ni-affinity chromatography](image)

**Figure 6.14:** SDS-PAGE and western blot analysis of *Glycine max* IFS fused to *Rhodococcus* sp. RhF reductase after purification with Ni-affinity chromatography. A: SDS PAGE and B: western blot of the purification of IFS-RhF (1 = soluble fraction after sonication, 2 = insoluble fraction after sonication, 3 = wash step, 4 = elution of unspecific protein, 5 = eluted IFS-RhF).

6.4.2 CYP73A5tr fusions

Resting cell assays were conducted on 73A5tr-lamATR2tr, 73A5tr-licATR2tr and 73A5tr-RhF expressed in *E. coli* Rosetta 2 (DE3). All fusions converted cinnamic acid to coumaric acid, with 73A5tr-lamATR2tr and 73A5tr-licATR2tr producing 33% and 23% after 24 h, respectively (Figure 6.15). The plant bacterial fusion 73A5tr-RhF shown with 3.5% genistein production a much lower conversion after 24 h. There were no decrease in cinnamic acid, nor increase in coumaric acid in the cells harbouring the empty vector only (data not shown).
Figure 6.15: Cinnamate-4-hydroxylase activity of Arabidopsis P450 73A5tr fused to Arabidopsis ATR2tr in a resting cell assay

73A5tr = truncated version of cinnamate-4-hydroxylase, RhF = P450 reductase domain from *Rhodococcus* sp., ATR2tr = truncated version of P450 reductase from Arabidopsis ATR2. The error bars represent the mean of four independent replicas ± standard deviation.
Subsequent, SDS PAGE as well as a western blot analyses were performed for the soluble and the insoluble fraction of the expressed 73A5tr fusions (Figure 6.16 and Figure 6.17). No bands at the predicted size of the fusion enzymes were found in the soluble protein fraction. However bands at the predicted size for all fusions (73A5tr-RhF = 91 kDa, 73A5tr-lamATR2tr = 127 kDa, 73A5tr-licATR2tr = 127 kDa) and the RhF alone (36 kDa) were detected in the insoluble fraction. No band was visible for the negative Rosetta 2 cell control.

Figure 6.16: SDS-PAGE and western blot analysis of Arabidopsis P450 73A5tr fused Arabidopsis ATR2tr or to Rhodococcus sp. RhF reductase in the soluble protein fraction
A SDS PAGE and B western blot analysis against anti His-antibodies (1 = 73A5tr-lamATR2tr, 2 = 73A5tr-licATR2tr 3 = 73A5tr-RhF, 4 = lamATR2tr control, 5 = RhF control, 6 = Rosetta 2 cells)
6.5 Discussion

The IFS-CPR fusion, an artificial fusion of IFS from *G. max* coupled to the 71 amino acid truncated reductase CPR from *C. roseus*, was developed by Prof. Koffas’s research group. The genes were fused using an artificial linker λ (GST) to reduce the formation of secondary structure and different modifications of the IFS N-terminus were tested to try to increase activity as measured by the production of genistein from naringenin. In their studies, using JM109 cells, the highest yield with around 15 μM (30%) genistein was seen when the construct contained the IFS with a synthetic mammalian peptide at the N-terminus fused to the 71 amino acid truncated CPR (*C. roseus*) through the linker λ. The published work was repeated here using the same cell strain and experimental methods, and the fusion protein tested: just 5 μM (10%) genistein was produced after 24 h. Additionally, the Rosetta 2 (DE3) strain, which contained an additional plasmid for rare codons, was tested. Five times more genistein (25 μM = 50%) in vivo was produced by this strain after 24 h and thus it was chosen for further analysis.

The plant P450 IFS was fused to the bacterial reductases RhF from *Rhodococcus* sp. creating the first plant-bacterial self-sufficient system. The RhF
reductase is part of a native P450-RhF fusion and was already used for engineering artificial fusions with various prokaryotic P450 such as P450cam and P450-XpIA heme.\textsuperscript{146}

A conversion of 3\% genistein was initially detected for IFS-RhF. This was optimized to almost 50\% genistein production by simplifying the protocol to an expression in M9 minimal medium instead of changing the medium after 3 h pre expression in LB to M9. Under these optimised conditions, IFS-CPR made 12 \(\mu\)M genistein, which is 15\% less than IFS-CPR in Rosetta 2 (DE3) under the conditions reported by Leonard and Koffas 2007\textsuperscript{306}.

Other groups tested the artificial plant fusion of the isoflavone synthase from red clover (\textit{Trifolium pratense}) (rcIFS) fused to rice (\textit{Oryza sativa}) reductase (rcIFS-riceCPR) or the reductases from \textit{C. roseus} (rcIFS-CPR). While the experimental conditions were not identical to those used here, making comparisons difficult, the rcIFS-riceCPR fusion showed 75\% conversion to genistein when expressed in LB after 12 h in resting cell assays, whereas only 26\% conversion was detected by using M9 medium.\textsuperscript{307}

In this project, two different expression temperatures (20 °C and 15 °C) were tested for IFS-CPR and IFS-RhF. The conversion of naringenin to genistein was reduced about 22\% and 30\% for IFS-CPR and IFS-RhF, respectively by dropping the temperature from 20 °C to 15 °C.

Kim \textit{et al.} reported the conversion of naringenin to genistein by rcIFS-riceCPR, performed at 28 °C in a growing cell assay with 63\% after 12 h,\textsuperscript{307} which was higher than by IFS-RhF (5\% conversion to genistein after 12 h at 20 °C, this project), possibly caused by the lower temperature used in the growing cell assay. The red clover IFS expressed in yeast produced 18\% genistein, which was increased to 88\% genistein when co expressed with the riceCPR in \textit{S. cerevisiae}.\textsuperscript{442} The endogeneous reductase of \textit{S. cerevisiae} did not transfer the electrons from the cofactor to the P450 rcIFS as effectively as the plant reductase riceCPR.

While the isoflavone synthase activity was relatively easy to measure, detecting expressed IFS protein by SDS-PAGE or western blot analysis was difficult as these methods are less sensitive.

Unfortunately, no signal was detected for IFS-CPR by western blot analysis (anti-His-antibody), possibly through the protein fold not exposing the His-tag or
because this method is less sensitive than the enzyme assay and the level of IFS was below the limits of detection using this technique. Leonard and Koffas reported in their publication the in vivo activity of IFS-CPR without recognising any in vitro activity of purified enzyme or the detection of the fusion by SDS PAGE or western blot. The fusion IFS-RhF was detected after N-affinity chromatography by western blot analysis using His-antibodies. A stronger signal for the fusion IFS-RhF was received when blotted against specific RhF reductase antibodies. IFS fused to the Arabidopsis reductase ATR2tr (IFS-lamATR2tr and IFS-licATR2tr) was only detected in the insoluble protein fraction by western blot analysis with specific antibodies raised against the reductase ATR2tr, suggesting membrane associated fusion protein. Additional degradation of the fusion to IFS and ATR2tr was found.

The Arabidopsis CYP73A5 (cinnamate-4-hydroxylase), lacking 26 amino acids at its N-terminus of the hydrophobic membrane anchor, was fused to the truncated version of the Arabidopsis ATR2tr and to the RhF reductase from Rhodococcus sp. to compare their activities. Both fusions showed the ability to hydroxylate cinnamic acid to coumaric acid in resting cell assays. The 73A5tr fused to ATR2tr resulted showed 20% higher activity than 73A5tr-RhF (3.5% coumaric acid after 24 h) demonstrating in this case, that the coupling of a P450 and reductase originating from the same plant are more efficient than when the two enzymes are derived from different organisms. This is in agreement with the observation also made for the flavonoid-3’,5’-hydroxylase (CYP75) from C. roseus fused to the CPR of C. roseus, which had a two-fold higher specific acitivity than the hydroxylase CYP75 from C. roseus fused to Petunia hybrid reductase. However, in conflict with this, studies by Mizutani found that the turnover number (amount of substrate molecules that the P450 can convert to product per catalytic site and per unit of time) of recombinant CYP73A5 expressed in insect cells was similar for the conversion of cinnamic acid to coumaric acid in combination with ATR2, or the reductase from Vigna radiate (mung bean) in-vitro. Urban et al. detected a 4-fold higher turnover rate for cinnamate-4-hydroxylase from barley (Helianthus tuberosus) co-expressed with yeast (S. cerevisiae) reductase in yeast microsomes. This research group compared also the Arabidopsis CYP73A4 co-expressed in yeast with a yeast reductase,
Arabidopsis ATR1 and ATR2 (yeast strains WAT11 and WAT21, respectively) and found that the turnover rate of CYP73A5 with yeast reductase was two-fold higher than with either of the two Arabidopsis reductases.\textsuperscript{166}

Additionally, two different linkers (lam- and lic-linker) connecting the P450 73A5tr and the reductase ATR2tr were tested and the activity of the fusion containing the lam-linker was 10\% higher than with the lic-linker. Similar was seen for the IFS fusions, where IFSsyn-lamATR2tr produced two-fold more genestein than IFSsyn-licATR2tr. Both linkers were dependent on the growth medium (Chapter 5, page 137). Modifications at the N-terminus, functioning as membrane anchor, and the linker region joining P450 and reductase may increase expression and activity.

The reductases ATR2tr and RhF were expressed independently and in a soluble and active form in \textit{E. coli} Rosetta 2 (DE3). The P450s, lacking their hydrophobic membrane anchor, fused to the reductases (ATR2tr or RhF) were expressed as active enzymes; however, they still appear to be associated to the membrane and are found in the insoluble protein and membrane fraction. This finding implies even so the membrane anchor has been removed that the P450 may still be associated to the membrane. 73A5tr-fusions lacking the membrane anchor were still present in the insoluble fraction. The solubility may be increased by modification of the P450 N-terminus, for example replacing the N-terminal region with a signal sequence, such as the endogeneous \textit{E. coli} OmpA signal sequence\textsuperscript{443} or the sequence of the bovine steroid 17\(\alpha\)-hydroxylase (CYP17A), which would also be useful for \textit{in vitro} assays\textsuperscript{318,444}.

Previous studies reported that plant P450s fused to a plant reductase have been actively expressed in \textit{E. coli}. Here, optimisation studies have shown that the expression strain is important and expression in the Rosetta 2 strain resulted in the highest activity in resting cell assays. Additional codon optimisation for \textit{E. coli} expression of IFS increased the activity more than 10\% as well as the linker region connecting P450 and reductase. Future further optimisation areas to test would include the linker region and also the N-terminus of the P450.
Soluble expression of plant P450s has been notoriously difficult and proven a bottleneck in characterisation of these enzymes. This platform technology shows promise for the study of plant P450s using artificial fusions with plant or bacterial reductases, because of the advantages of expressing a single fusion enzyme instead of two separate ones. Additionally, the platform could be used for *in vivo* activity tests screening a range of substrates, or conversely, to screen a large number of P450s for activity to a specific product of interest with the aim of identifying and characterising new P450 functions.
7 Final Discussion

Biotransformation is a fast growing sector in industry and is more and more displacing synthetic chemistry. The benefits of using biocatalysts for transformation include the ability to make enantio- and regiospecific products as well as frequently having more environmentally friendly production conditions when compared to traditional chemistries.\textsuperscript{445,446}  

Cytochromes P450 are fascinating due to their occurrence in all kingdoms of life as well as their broad substrate range, regio- and stereospecificity. P450s require a corresponding reductase and the cofactor NAD(P)H for their reaction. Human P450s occur predominantly bound to the endoplasmic reticulum of liver cells, where they are involved in the metabolism of drugs and medicine, hence the activities of P450s are intensively studied. Bacterial P450 systems are cytosolic and have also been the subject of much research due to their solubility, as well as their roles in secondary metabolism. Plants possess significantly more P450s than found in mammals or bacteria, with often more than 200 P450s in a single plant species. Plant P450s are mainly involved in the reactions of secondary metabolism, and are stimulated under stress conditions. They are insoluble due to a strong association with an endoplasmic reticulum or organelle membrane, for example the membrane of the chloroplast by a serine-threonine rich sequence\textsuperscript{351}, and are therefore not easy to express actively as recombinant proteins. As a result the plant P450s have been examined in much less detail than their counterparts in other kingdoms and so for most plant P450s a function is still not known. For example, a search of publications on Arabidopsis P450s (reviewed in Chapter 1) showed that for 75% of Arabidopsis P450s the native substrate has not yet been identified. Given these limitations, the development of a specific system to express and assay active plant P450s is necessary.

A platform technology containing plant P450s fused to appropriate reductases was engineered in this project. This technology will provide a system for the study of artificial plant P450-reductase fusions and investigation of P450s with unknown function with the facility to readily clone P450s in frame, upstream of a reductase. Two great advantages of using fusion proteins are that just one protein needs to be prepared instead of two separate enzymes, and there is the potential to improve activity by more efficient electron transfer from the reductase to the P450. In this
project, attention was concentrated on P450s from Arabidopsis due to the availability of annotated gene sequences and that this species is of interest many members of the plant science community. Native Arabidopsis reductases were tested to optimise activity of the P450 fusions. Encoded in the Arabidopsis genome are three Arabidopsis P450 reductases: ATR1, ATR2 and ATR3. While ATR3 is involved primarily in embryo development, ATR1 and ATR2 are expressed predominantly throughout the spatial and temporal development of the plant. Both proteins, along with their native membrane anchor regions, have been expressed and characterised as membrane proteins in yeast and insect cells. Although, the main focus of this project is the rapid screening recombinant P450s for activity in vivo, expression of soluble fusions would be advantageous for more detailed biochemical characterisation. Hence, truncated forms of the reductases (ATR1tr and ATR2tr) were expressed without N-terminal hydrophobic membrane anchors and cloned into pET-YSBLIC. Expression in *E. coli* and purification by nickel-affinity chromatography were optimised to attain soluble, active ATR1tr and ATR2tr. Activity assays, using cytochrome c, showed that ATR2tr was 20 times more active than ATR1tr. ATR2tr was therefore selected for creating the fusion platform technology and cloned into pET-YSBLIC using traditional, non-LIC, cloning methods. A linker sequence was also inserted at the 5’ end of ATR2tr to connect the encoded P450 to the reductase. The lam-linker was developed based on the tripeptide linker GST which is often used for artificial plant P450 fusion systems. A second linker, called lic-linker, which has been successfully used for fusions of prokaryotic P450s with the RhF reductase from *Rhodococcus*, was used. In addition to the ATR2tr, the RhF reductase was cloned and tested for activity in *E. coli* with plant P450s to create plant-bacterial P450 fusions. Four different P450s (CYP93C, CYP73A5, CYP81D8 and CYP82E4) were selected to show proof-of-principle of the novel plant fusion system (Table 7.1).
Table 7.1: List of P450-reductase fusion constructs created in this project and their tested substrates

<table>
<thead>
<tr>
<th>P450</th>
<th>activity</th>
<th>Conversion after 24 h in resting cell assays</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>substrate</td>
<td>product</td>
</tr>
<tr>
<td>IFS (CYP93C)</td>
<td>naringenin</td>
<td>genistein</td>
</tr>
<tr>
<td>IFSopt (CYP93C)</td>
<td>naringenin</td>
<td>genistein</td>
</tr>
<tr>
<td>73A5tr (opt.)</td>
<td>cinnamic acid</td>
<td>coumaric acid</td>
</tr>
<tr>
<td>CYP82E4</td>
<td>nicotine</td>
<td>nornicotine</td>
</tr>
<tr>
<td>81D8tr</td>
<td>TNT, DNT, MTS, EC</td>
<td>not tested</td>
</tr>
</tbody>
</table>

opt = gene codon optimized for E. coli expression, TNT = 2,4,6-trinitrotoluene, DNT = 2,4-dinitrotoluene or 2,6-dinitrotoluene, MTS = methyl-tolyl-sulphide, EC = 7-ethoxycoumarin, * experiment performed as growing cell assay

Comparison of reductases, codon usage and linker regions for fusion expression

A comparison of isoflavone synthase (IFS) (CYP93C1 from Glycine max) fused to different reductases was performed to test that the P450 activity was dependent on the electron transfer by the reductase. As shown in Table 7.1, IFS-CPR and IFS-RhF resulted in the highest activities. While the activity of the nativeIFS-ATR2tr fusion was improved following codon optimisation for expression in E. coli, activity was still less than for the IFS-CPR and IFS-RhF fusions.

A similar comparison using the codon optimised Arabidopsis P450 cinnamate-4-hydroxylase without the hydrophobic membrane anchor (73A5tr) revealed that, as shown in Table 7.1, 73A5tr-ATR2tr had higher activity than the CYP73A5-bacterial fusion (73A5tr-RhF).

Expression of native CYP82E4, N-demethylase from Nicotiana tabacum and native Arabidopsis CYP81D8 (of unknown function) fused to ATR2tr did not result in activity using the tested substrates.

To assess linker sequences joining the P450 to the ATR2tr the lam- and lic-linkers were compared. Although results varied according to experimental conditions, overall the lam-linker gave higher activity.

The results show that plant P450s work best when fused to a reductase from the same species. In the absence of a species-specific reductase, the bacterial RhF reductase from Rhodococcus could be used as a universal reductase. RhF is part of
a native bacterial fusion which has evolved to live in an environment alongside
the P450. Although in the native fusion the P450 is fused to the C-terminus of
RhF reductase (Figure 1.11B), in the studies here the plant P450 was fused to the
N-terminus of the RhF reductase (Figure 7.1). The results also show that codon
optimisation for expression in *E. coli* improves activity. As the costs in gene
synthesis have dropped significantly in recent years, along with improvements in
software tools to optimise for *E. coli* codon bias and additional modifications,
discussed below, at a large (100-1000 sequences) scale, this could now be a viable
step in enhancing performance of the screen.

The plant P450-plant reductase fusions, or if the plant-specific reductase is
unavailable, plantP450-RhF fusions, could be used as artificial fusion systems for
preliminary, high-throughput substrate screens. A secondary screen for P450s not
expressed during the preliminary screen, could be performed whereby activity is
achieved by using codon optimisation for *E. coli* expression and alternative
reductases. These procedures could also be applied to increase the expression and
activity of promising P450s for more detailed characterisation.

**Improving performance in cell assays**

It was observed that the maximum conversion of substrate into product was 40 %
(seen with cinnamic acid to coumaric acid after 48 hours). Given that substrate
levels indicated no other products were being formed, this means that the assay is
less sensitive than it could be, and potentially expensive, substrate is wasted.
Analysis of the data (Figure 5.13) showed that the conversion was non-linear,
with activity slowing down over time. Possible reasons for this are that the
NADPH in the cells has been depleted to a limiting concentration. NADPH is
produced via the pentose phosphate pathway. Adding glucose, as a supplemental
carbon source, to the sample during the assay could address this possible
limitation in substrate conversion in future experiments. Following Michaelis-
Menten kinetics, at limiting substrate concentrations, enzyme rates decrease with
substrate concentration decrease. The amount of conversion, and therefore
sensitivity, could be increased by the addition of more substrate. Ensuring that a
non-limiting substrate concentration is used, activity with increasing substrate
concentration could be monitored.
Enhancing solubility of the fusions

All the reductases tested in this project were expressed as active, soluble forms in *E. coli*. However, while fusion proteins were found in both soluble and insoluble protein fractions, activity was only found in the insoluble fractions. These results indicate that active fusions are associated with a membrane even though the membrane targeting signal was removed during the cloning, and that this membrane association is P450-specific. Literature reported hydrophobic regions on the surface of the P450 like the F-G loop, which is also part of the substrate access channel, show interaction with membranes. Studying existing P450 structures, as well as analysing each target P450 amino acid sequence (used for the fusion), would show hydrophobic parts of the enzyme. These could be modified if the gene is re-synthesised for codon optimisation, as discussed above, to improve solubility of the whole fusion protein. Optimising the N-terminal truncation of the P450s further could also lead to improvements in solubility and expression yield. Studying the hydrophobic regions of the P450s responsible for the insolubility by interacting with the membrane could identify regions to be modified. Beyond this, the mechanism of the electron transfer between reductase and P450 could be studied, with particular attention to the linker region, with the aim of improving efficiency.

Future directions for P450-reductase fusion systems

In this project two fusions systems containing plant P450s fused to the N-terminus of the Arabidopsis reductase ATR2tr (Figure 7.1A) or the *Rhodococcus* RhF (Figure 7.1B) reductase have been characterised. Overall, it was shown that both fusion systems resulted in active protein and the activity was dependent on the P450. This is the first reported artificial plant P450-bacterial reductase fusion. Even though the bacterial and plant reductases are structurally divergent, it is remarkable that they both support the P450 reaction.
Figure 7.1: Schematic organisation of artificial A) plant-plant fusions (P450-ATR2tr) and B) plant bacterial (P450-RhF) fusions
Both fusions are lacking the hydrophobic membrane anchor at the N-terminus of the P450s and are fused C-terminally to the reductase. Due to hydrophobic regions on the surface of the P450 the fusion might be still membrane associated.

The main outcome of this project is a platform technology that is a useful tool for the investigation of novel plant P450 activities. It can be employed for high-throughput cloning via a ligation independent cloning system enabling a large number of P450s to be easily cloned, expressed as fusions in E. coli and screened using resting cell assays. However, there are still deficiencies to this system which can be addressed. In this way, this technology presents a new opportunity to search libraries of plant P450 fusions against a wide range of substrates to discover new activities and new biocatalysts for the industrial future.
# Appendix A

## Table A1: Cytochromes P450 in Arabidopsis

<table>
<thead>
<tr>
<th>P450 (gene locus)</th>
<th>reaction</th>
<th>pathway</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>51G1 (At1g11680)</td>
<td>obtusifoliol 14α-demethylase</td>
<td>sterols/steroids</td>
<td>183,184</td>
</tr>
<tr>
<td>51G2 (At2g17330)</td>
<td>obtusifoliol 14α-demethylase (pseudogene)</td>
<td></td>
<td>184</td>
</tr>
<tr>
<td>71A12 (At2g30750)</td>
<td>hydroxylation and N-demethylation of pyrazoxyfen</td>
<td>metabolism of herbicide pyrazoxyfen, camalexin biosynthesis in roots</td>
<td>185,186</td>
</tr>
<tr>
<td>71A13 (At2g307700)</td>
<td>dehydration of indole acetaldoxime to indole-3-acetonitrile</td>
<td>camalexin biosynthesis</td>
<td>187</td>
</tr>
<tr>
<td>71A14 (At5g24960)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A15 (At5g24950)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A16 (At5g42590)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A17P</td>
<td>pseudogene fragment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A18 (At1g11610)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A19 (At4g13290)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A20 (At4g13310)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A21 (At3g48320)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A22 (At3g48310)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A23 (At3g48300)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A24 (At3g48290)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A25 (At3g48280)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A26 (At3g48270)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A27P (At4g20240)</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71A28 (At4g20235)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B2 (At1g13080)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B3 (At3g26220)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B4 (At3g26280)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B5 (At3g53280)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B6 (At2g24180)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B7 (At1g13110)</td>
<td>deethylation of 7-ethoxycoumarin (with cumene hydroperoxide as electron donor)</td>
<td></td>
<td>188</td>
</tr>
<tr>
<td>71B8 (At5g35715)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B9 (At2g02580)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B10 (At5g57260)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B11 (At5g25120)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B12 (At5g25130)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B13 (At5g25140)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B14 (At5g25180)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B15 (At3g26830)</td>
<td>conversion of s-dihydrocamalexic acid to camalexin</td>
<td>camalexin</td>
<td>189,190</td>
</tr>
<tr>
<td>71A16 (At3g26150)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B17 (At3g26160)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B18P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P450 (gene locus)</td>
<td>reaction</td>
<td>pathway</td>
<td>references</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>71B19 (At3g26170)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B20 (At3g26180)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B21 (At3g26190)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B22 (At3g26200)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B23 (At3g26210)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B24 (At3g26230)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B25 (At3g26270)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B26 (At3g26290)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B27 (At1g13070)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B28 (At1g13090)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B29 (At1g13100)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B30P (At3g53290)</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B31 (At3g53300)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B32 (At3g53305)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B33</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B34 (At3g26300)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B35 (At3g26310)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B36 (At3g26320)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B37 (At3g26330)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71B38 (At3g44250)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72A7 (At3g14610)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72A8 (At3g14620)</td>
<td>No native function known, hydroxylation of quinidine</td>
<td>quinidine metabolism</td>
<td>341</td>
</tr>
<tr>
<td>72A9 (At3g14630)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72A10 (At3g14640)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72A11 (At3g14650)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72A12P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72A13 (At3g14660)</td>
<td>putative P450</td>
<td>upregulated in response to cis-jasmonate</td>
<td>377</td>
</tr>
<tr>
<td>72A14 (At3g14680)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72A15 (At3g14690)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72C1 (At1g17060)</td>
<td>exact substrate not identified (binds typhasterol)</td>
<td>degradation of brassinosteroids</td>
<td>293,448-450</td>
</tr>
<tr>
<td>73A5 (At2g30490)</td>
<td>4-hydroxylation of t-cinnamic acid to p-coumaric acid; hydroxylation of cinnamic acid analogs</td>
<td>phenylpropanoid pathway, lignin biosynthesis</td>
<td>166,191,192</td>
</tr>
<tr>
<td>74A1 (At5g42650)</td>
<td>allene oxide synthase for linoleic acid hydroperoxide and linolenic acid hydroperoxide</td>
<td>jasmonate acid biosynthesis</td>
<td>37,45,193-196</td>
</tr>
<tr>
<td>74B2 (At4g15440)</td>
<td>hydroperoxide lyase for linoleic acid hydroperoxide and linolenic acid hydroperoxide</td>
<td>oxylipin pathway, jasmonate biosynthesis</td>
<td>197-199</td>
</tr>
<tr>
<td>75B1 (At5g07990)</td>
<td>3'-hydroxylation of narigenin and dihydrokaempferol</td>
<td>phenylpropanoid pathway, flavonoid biosynthesis</td>
<td>200</td>
</tr>
<tr>
<td>76C1 (At2g45560)</td>
<td>10-hydroxylation of geraniol</td>
<td>terpenoid indole alkaloid biosynthesis</td>
<td>201</td>
</tr>
<tr>
<td>76C2 (At2g45570)</td>
<td>putative P450</td>
<td>possibly involved in cell death</td>
<td>451</td>
</tr>
<tr>
<td>P450 (gene locus)</td>
<td>reaction</td>
<td>pathway</td>
<td>references</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>76C3 (At2g45580)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76C4 (At2g45550)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76C5 (At1g33730)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76C6 (At1g33720)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76C7 (At3g61040)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76C8P (At3g61035)</td>
<td>pseudogene fragment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76G1 (At3g52970)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77A4 (At1g506460)</td>
<td>epoxidation of unsaturated C18 fatty acids (oleic, linoleic, α-linolenic acid)</td>
<td>fatty acid metabolism</td>
<td>178</td>
</tr>
<tr>
<td>77A5P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77A6 (At3g10570)</td>
<td>in-chain 10-hydroxylation of 16-hydroxyphalmitate</td>
<td>cutin biosynthesis</td>
<td>202</td>
</tr>
<tr>
<td>77A7 (At3g10560)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77A8P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77A9 (At5g04630)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77B1 (At1g11600)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78A5 (At1g13710)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78A6 (At2g46660)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78A7 (At5g09970)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78A8 (At1g01190)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78A9 (At3g61880)</td>
<td>putative P450</td>
<td>possibly involved in fruit development</td>
<td>452</td>
</tr>
<tr>
<td>78A10 (At1g74110)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79A2 (At5g05260)</td>
<td>conversion of phenylalanine to oxime</td>
<td>benzylglucosinolate biosynthesis</td>
<td>203</td>
</tr>
<tr>
<td>79A3P (At5g35917)</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79A4P (At5g35920)</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79B2 (At4g39950)</td>
<td>conversion of tryptophan and tryptophan analogs to oxime</td>
<td>indole glucosinolate biosynthesis, camalexin biosynthesis, auxin biosynthesis</td>
<td>204–209</td>
</tr>
<tr>
<td>79B3 (At2g22330)</td>
<td>conversion of tryptophan and tryptophan analogs to oxime</td>
<td>indole glucosinolate biosynthesis, camalexin biosynthesis, auxin biosynthesis</td>
<td>204–209</td>
</tr>
<tr>
<td>79B4P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79C1 (At1g79370)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79C2 (At1g58260)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79C3P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79C4P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79C5P (At1g58265)</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79F1 (At1g16410)</td>
<td>N-hydroxylation of homo- to tetrahomo-methionine (n = 3 to 6) (short-chain methionine derivatives to their aldoximes)</td>
<td>biosynthesis of aliphatic glucosinolates</td>
<td>210-213</td>
</tr>
<tr>
<td>P450 (gene locus)</td>
<td>reaction</td>
<td>pathway</td>
<td>references</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>79F2 (At1g16400)</td>
<td>N-hydroxylation of long chain penta and hexahomomethionine to their aldoxime</td>
<td>biosynthesis of aliphatic glucosinolates</td>
<td>211-213</td>
</tr>
<tr>
<td>81D1 (At5g36220)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81D2 (At4g37360)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81D3 (At4g37340)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81D4 (At4g37330)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81D5 (At4g37320)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81D6 (At2g23220)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81D7 (At2g23190)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81D8 (At4g37370)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81D10 (At1g66540)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81D11 (At3g28740)</td>
<td>putative P450</td>
<td>induced by cis-jasmone</td>
<td>377,378</td>
</tr>
<tr>
<td>81F1 (At4g37430)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81F2 (At5g57220)</td>
<td>4-hydroxylation of indole-3-ylmethyl to 4-hydroxyindole-3-ylmethyl glucosinolate</td>
<td>indole glucosinolate biosynthesis</td>
<td>214-216</td>
</tr>
<tr>
<td>81F3 (At4g37400)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81F4 (At4g37410)</td>
<td>putative P450</td>
<td>possibly involved in the indole glucosinolate biosynthesis</td>
<td>217</td>
</tr>
<tr>
<td>81G1 (At5g67310)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81H1 (At4g37310)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81K1 (At5g10610)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81K2 (At5g10600)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82C2 (At4g31970)</td>
<td>5-hydroxylation of 8-methoxypsoralen to 5-hydroxy-8-methoxypsoralen</td>
<td>metabolism of tryptophan-derived secondary metabolites, possibly involved in jasmonic acid induced indole glucosinolates</td>
<td>218,219</td>
</tr>
<tr>
<td>82C3 (At4g31950)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82C4 (At4g31940)</td>
<td>5-hydroxylation of 8-methoxypsoralen to 5-hydroxy-8-methoxypsoralen</td>
<td>Fe deficiency response, possibly through an IDE1-like mediated pathway</td>
<td>218,220</td>
</tr>
<tr>
<td>82F1 (At2g25160)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82G1 (At3g25180)</td>
<td>(E,E)-geranylinalool and the sesquiterpenoid (E)-nerolidol into the acyclic volatile C_{14}-homoterpene 4,8, 12-trimethylthdeca-1,3,7, 11-tetraene (TMTT) and the C_{13}-homoterpenes 4,8-dimethyl-1,3,7-nonathene (DMNT), respectively</td>
<td></td>
<td>221</td>
</tr>
<tr>
<td>P450 (gene locus)</td>
<td>reaction</td>
<td>pathway</td>
<td>references</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>83A1 (At4g13770)</td>
<td>oxidation of methionine-derived oximes oxidation of p-hydroxyphenylacetaldoxime, indole-3-acetaldoxime, conversion of aldoximes to thiohydroximates</td>
<td>biosynthesis of aliphatic glucosinolate</td>
<td>222-223</td>
</tr>
<tr>
<td>83B1 (At4g31500)</td>
<td>oxidation of indole-3-acetaldoxime</td>
<td>biosynthesis of indole glucosinolate</td>
<td>222,223,225</td>
</tr>
<tr>
<td>84A1 (At4g36220)</td>
<td>5-hydroxylation of coniferaldehyde, coniferyl alcohol and ferulic acid</td>
<td>phenylpropanoid pathway, biosynthesis of lignin</td>
<td>167,226,227</td>
</tr>
<tr>
<td>84A4 (At5g04330)</td>
<td>putative ferulate-5-hydroxylase</td>
<td></td>
<td>228,229</td>
</tr>
<tr>
<td>85A1 (At5g38970)</td>
<td>C6-oxidase for 6-deoxyxycastasterone to castasterone, other steroids conversion of castasterone to brassinolide</td>
<td>biosynthesis of brassinolide</td>
<td>230-234</td>
</tr>
<tr>
<td>85A2 (At3g30180)</td>
<td>C6-oxidase for 6-deoxyxycastasterone to castasterone, other steroids conversion of castasterone to brassinolide</td>
<td>biosynthesis of brassinolide</td>
<td>230-233,235,236</td>
</tr>
<tr>
<td>86A1 (At5g58860)</td>
<td>ω-hydroxylation of saturated and unsaturated C12 to C20 fatty acids</td>
<td>fatty acid metabolism, biosynthesis of cutin, biosynthesis of suberin</td>
<td>237-240</td>
</tr>
<tr>
<td>86A2 (At4g00360)</td>
<td>ω-hydroxylation of saturated and unsaturated C12 to C18 fatty acids</td>
<td>fatty acid metabolism, biosynthesis of cutin</td>
<td>238,241,242</td>
</tr>
<tr>
<td>86A4 (At1g01600)</td>
<td>ω-hydroxylation of saturated and unsaturated C12 to C18 fatty acids</td>
<td>fatty acid metabolism, biosynthesis of cutin</td>
<td>202,238,242,243</td>
</tr>
<tr>
<td>86A7 (At1g63710)</td>
<td>ω-hydroxylation of lauric acid</td>
<td>fatty acid metabolism, biosynthesis of cutin</td>
<td>238,242</td>
</tr>
<tr>
<td>86A8 (At2g45970)</td>
<td>ω-hydroxylation of saturated and unsaturated C12 to C18 fatty acids</td>
<td>fatty acid metabolism, biosynthesis of cutin</td>
<td>238,242,244</td>
</tr>
<tr>
<td>86B1 (At5g23190)</td>
<td>ω-hydroxylation for long chain fatty acid (C22 and C24)</td>
<td>biosynthesis of suberin, fatty acid metabolism, polyester monomer biosynthesis</td>
<td>38,245</td>
</tr>
<tr>
<td>86B2 (At5g08250)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>86C1 (At1g24540)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>86C2 (At3g26125)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>86C3 (At1g13140)</td>
<td>hydroxylation of fatty acids (C12, C14, C14:1, C16)</td>
<td>fatty acid metabolism</td>
<td>246</td>
</tr>
<tr>
<td>86C4 (At1g13150)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>87A2 (At1g12740)</td>
<td>putative P450</td>
<td>possibly involved in auxin signaling</td>
<td>453</td>
</tr>
<tr>
<td>87A3P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P450 (gene locus)</td>
<td>reaction</td>
<td>pathway</td>
<td>references</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>88A3 (At1g05160)</td>
<td>oxidation of ent-kaurenoic acid to gibberillin A12 in three steps</td>
<td>biosynthesis of gibberellins</td>
<td>39,247</td>
</tr>
<tr>
<td>88A4 (At2g32440)</td>
<td>oxidation of ent-kaurenoic acid to gibberillin A12 in three steps</td>
<td>biosynthesis of gibberellins</td>
<td>39,247</td>
</tr>
<tr>
<td>89A2 (At1g64900)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89A3 (At5g61320)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89A4 (At2g12190)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89A5 (At1g64950)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89A6 (At1g64940)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89A7 (At1g64930)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89A9 (At3g03470)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90A1 (At5g05690)</td>
<td>23α-hydroxylation of sterols</td>
<td>biosynthesis of brassinolide</td>
<td>171,172,248,249</td>
</tr>
<tr>
<td>90B1 (At3g50660)</td>
<td>22α-hydroxylation of C27, C28, C29 sterols</td>
<td>biosynthesis of brassinolide</td>
<td>169,249-251</td>
</tr>
<tr>
<td>90C1 (At4g36380)</td>
<td>C23-hydroxylation of sterols (typhasterol to castasterone)</td>
<td>biosynthesis of brassinolide</td>
<td>252-255</td>
</tr>
<tr>
<td>90D1 (At3g13730)</td>
<td>C23-hydroxylation of sterols</td>
<td>biosynthesis of brassinolide</td>
<td>253-255</td>
</tr>
<tr>
<td>93D1 (At5g06900)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>94B1 (At3g63450)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>94B2 (At3g01900)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>94B3 (At3g48520)</td>
<td>12-hydroxylation of jasmonoyl-L-isoleucine</td>
<td>oxidative catabolism of jasmonate, jasmonate mediated signaling pathway</td>
<td>256,257</td>
</tr>
<tr>
<td>94C1 (At2g27690)</td>
<td>ω-hydroxylation of saturated and unsaturated C12 to C18 fatty acids and hydroxylation of ω-hydroxy fatty acid into dicarboxylic fatty acid</td>
<td>fatty acid metabolism</td>
<td>240,258</td>
</tr>
<tr>
<td>94D1 (At1g34540)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>94D2 (At3g56630)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>94D3P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A1 (At2g23180)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A2 (At4g32170)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A3 (At1g65340)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A4 (At5g52320)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A5 (At2g21910)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A6P (At5g51900)</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A7 (At1g47630)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A8 (At1g47620)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A9 (At4g39480)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A10 (At4g39490)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A11 (At4g39500)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A12 (At4g39510)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P450 (gene locus)</td>
<td>reaction</td>
<td>pathway</td>
<td>references</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>96A13 (At5g02900)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A14P (At1g66030)</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96A15 (At1g577750)</td>
<td>hydroxylation of midchain alkane to alcohols and second hydroxylation to ketone</td>
<td>biosynthesis of wax</td>
<td>259</td>
</tr>
<tr>
<td>97A3 (At1g31800)</td>
<td>β-hydroxylation of carotene</td>
<td>carotenoid metabolism</td>
<td>260,261</td>
</tr>
<tr>
<td>97B3 (At4g15110)</td>
<td>β-hydroxylation of β-carotene to zeaxanthin via β-cryptoxanthin</td>
<td>carotenoid metabolism</td>
<td>262</td>
</tr>
<tr>
<td>97C1 (At3g53130)</td>
<td>ε-hydroxylation of β,ε-carotene</td>
<td>carotenoid metabolism</td>
<td>260,263</td>
</tr>
<tr>
<td>98A3 (At2g40890)</td>
<td>3-hydroxylation of p-coumarate to caffeic acid and 3-hydroxylation of coumaroyl-esters (shikimate and quinate esters)</td>
<td>phenylpropanoid pathway, biosynthesis of lignin monomers and soluble phenolics</td>
<td>168,264-267</td>
</tr>
<tr>
<td>98A8 (At1g74540)</td>
<td>meta-hydroxylation of the three triferuloylspermidine phenolic rings, oxygenation of resveratrol</td>
<td>alternative phenolic pathway, pollen development</td>
<td>268,269</td>
</tr>
<tr>
<td>98A9 (At1g74550)</td>
<td>meta-hydroxylation of the three triferuloylspermidine phenolic rings, oxygenation of resveratrol</td>
<td>alternative phenolic pathway, pollen development</td>
<td>268,269</td>
</tr>
<tr>
<td>701A3 (At5g25900)</td>
<td>oxidation of ent-kaurene to ent-kaurenoic acid in three steps</td>
<td>biosynthesis of gibberelin</td>
<td>39,270-273</td>
</tr>
<tr>
<td>702A1 (At1g65670)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>702A2 (At4g15300)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>702A3 (At4g15310)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>702A4P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>702A5 (At4g15393)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>702A6 (At4g15396)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>702A7P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>702A8 (At3g30290)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>703A2 (At1g01280)</td>
<td>in-chain monohydroxylation of saturated fatty acids (C10-C16)</td>
<td>biosynthesis of sporopollenin, pollen development</td>
<td>274</td>
</tr>
<tr>
<td>704A1 (At2g44890)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>704A2 (At2g45510)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>704B1 (At1g69500)</td>
<td>ω-hydroxylation of saturated, unsaturated and epoxy C16 and C18 fatty acids</td>
<td>biosynthesis of sporopollenin, pollen development</td>
<td>275,276</td>
</tr>
<tr>
<td>705A1 (At4g15330)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A2 (At4g15350)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A3 (At4g15360)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A4 (At4g15380)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A5 (At5g47990)</td>
<td>conversion of thaliandiol to desaturated thaliandiol</td>
<td>thalianol metabolic pathway</td>
<td>277</td>
</tr>
<tr>
<td>705A6 (At2g05180)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A8 (At2g27000)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P450 (gene locus)</td>
<td>reaction</td>
<td>pathway</td>
<td>references</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>705A9 (At2g27010)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A10P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A11P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A12 (At5g42580)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A13 (At2g14100)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A14P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A15 (At3g20080)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A16 (At3g20083)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A17P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A18 (At3g20090)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A19 (At3g20100)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A20 (At3g20110)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A21 (At3g20120)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A22 (At3g20130)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A23 (At3g20140)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A24 (At1g28430)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A25 (At1g50560)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A26P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A27 (At1g50520)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A28</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A29P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A30 (At3g20940)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A31P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A32 (At3g20950)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A33 (At3g20960)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>705A34P</td>
<td>pseudogene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>706A1 (At4g22690)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>706A2 (At4g22710)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>706A3 (At5g44620)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>706A4 (At4g12300)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>706A5 (At4g12310)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>706A6 (At4g12320)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>706A7 (At4g12330)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>707A1 (At4g19230)</td>
<td>8-hydroxylation of abscisic acid</td>
<td>abscisic acid catabolism</td>
<td>278-280</td>
</tr>
<tr>
<td>707A2 (At2g29090)</td>
<td>8-hydroxylation of abscisic acid</td>
<td>abscisic acid catabolism</td>
<td>278-281</td>
</tr>
<tr>
<td>707A3 (At5g45340)</td>
<td>8-hydroxylation of abscisic acid</td>
<td>abscisic acid catabolism</td>
<td>278,279,281-286</td>
</tr>
<tr>
<td>P450 (gene locus)</td>
<td>reaction</td>
<td>pathway</td>
<td>references</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>707A4 (At3g19270)</td>
<td>8-hydroxylation of abscisic acid</td>
<td>abscisic acid catabolism</td>
<td>278,279</td>
</tr>
<tr>
<td>708A1 (At1g55940)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>708A2 (At5g48000)</td>
<td>hydroxylation of thalianol to thaliandiol</td>
<td>thalianol metabolic pathway</td>
<td>277</td>
</tr>
<tr>
<td>708A3 (At1g78490)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>708A4 (At3g44970)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>709B1 (At2g46960)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>709B2 (At2g46950)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>709B3 (At4g27710)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>710A1 (At2g34500)</td>
<td>C22-desaturation of β-sitosterol to stigmasterol</td>
<td>biosynthesis of sterols</td>
<td>287,289</td>
</tr>
<tr>
<td>710A2 (At2g34490)</td>
<td>C22-desaturation of β-sitosterol and 24-epi-campesterol to stigmasterol and brassicasterol</td>
<td>biosynthesis of sterols</td>
<td>287</td>
</tr>
<tr>
<td>710A3 (At2g28850)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>710A4 (At2g28860)</td>
<td>C22-desaturation of β-sitosterol to stigmasterol</td>
<td>biosynthesis of sterols</td>
<td>288</td>
</tr>
<tr>
<td>711A1 (At2g26170)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>712A1 (At2g42250)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>712A2 (At5g06905)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>714A1 (At5g24910)</td>
<td>C13-hydroxylation of ent-kaurenoic acid to steviol</td>
<td>gibberillin catabolism</td>
<td>277,290</td>
</tr>
<tr>
<td>714A2 (At5g24900)</td>
<td>putative P450</td>
<td>gibberillin catabolism</td>
<td>290</td>
</tr>
<tr>
<td>715A1 (At5g52400)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>716A1 (At5g36110)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>716A2 (At5g36140)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>718 (At2g42850)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>720A1 (At1g73340)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>721A1 (At1g75130)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>722A1 (At1g19630)</td>
<td>putative P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>724A1 (At5g14400)</td>
<td>22α-hydroxylation of brassinosteroid</td>
<td>biosynthesis of brassinolide</td>
<td>291</td>
</tr>
<tr>
<td>734A1 (At2g26710)</td>
<td>initially classified as CYP72B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>735A1 (At5g38450)</td>
<td>trans-hydroxylation of isopentenyladenine, tri/di/monophosphates</td>
<td>biosynthesis of zeatin, cytokinin metabolism</td>
<td>294</td>
</tr>
<tr>
<td>735A2 (At1g67110)</td>
<td>trans-hydroxylation of isopentenyladenine, tri/di/monophosphates</td>
<td>biosynthesis of zeatin, cytokinin metabolism</td>
<td>294</td>
</tr>
</tbody>
</table>
Appendix B

B1 Gene sequences used in Chapter 3

CYP81D8 (from Arabidopsis) was isolated from plant cultivated in liquid culture. The truncated version of CYP81D8 (CYP81D8tr in italics) was created by deleting the hydrophobic N-terminal region.

ATGGAAACCAAAACCCTAATTTTCTCAATTCTTCTCGTTGTCTCTCTCTCATTT
ACTTAATTGGAAACTCAAGCGAAGCCAAATACACTTCCTCGAGTCCGGCAGATGGTC
ATTACCAGTATTGGTCATCTCCGGCTTTCTCAAACCACCGATTTCTCGACATTTC
CTCTCCCCCTCTCTCAATCCTAAAATATCATTCCGATCTTCTCCTCCCTCGACTCGGTA
ACCAGACTCTGGTTTCGTAACCTGTCACACTCGATCGAGGGAATGTTTCACCAA
AAACGACGTCGTACTGGGAACAGACCCAAACTTCATCCTCCTAAACACACCTGCG
TACGATTACACAAACCATTGATCGCAAGCTTCCTTACGGTACACTGGGCTAACCCTCC
GCCGCATCGGTCTCCTCGAGATATTTACTCCAATCCGGCTCTAAATAGCTGTCTTGC
TATTGGTTAAAAGAGTATCGAGACGCCTTACTGCTCAGGATGTTAGATGA
TCCGGATTGCTAAGCGAGATGGTAATGCTGACTGGAGGTTGAGGGTGGTGCTGCT
GCTGGAAACGCTTGTATTTACCTTCCGCTTTGTTTACGAGATTACGAGAAGATGGTAGA
CACGGGTTTAAAGACGCAAGCCTTGTGATACTGGGACAATCTGGGCTTTCGAGATGATG
TGAGAAACGAGAAGCTAAGGAGAAAGGAAACACTATGATCGATCACTTATGCTTTGCTTAC
CGACAGGAATCCCAACCGGATTATCTTCACGGGATCTCATATTAAAGAAAACATGC
TCGCTTTTGATATCTAGCCAGCAGACATCAGCGTTACGTAGATGGGACAT
GTCGAACTGTGTGTAACCACTCCGGATGTATGGAACAGGAGAGATGAAATCTG
AGAAGATAGTTTAGAGACGCTTATGGGATGATACTAGATATCTCAACCTGGCTT
ATCTCCAAAAACATTGTCTCAGTAAAGTTCCCTGCTTATCCTCGGCCTCTCATGCT
TCTTCCCTCAGTGTGCTTGGAGAAGATCGTAAAGTTGCAGGATACGCAATGCCCAGT
GCCACGATATATTGAGGCCAATCGTGTGCTATACACAGAGATCTCTGATG
ATGATCAATGACCTCAACCCGAGAGTTGAGAAAGAAGAAAGAGAACGCTTAC
GCTAATCGCTTGGTTAGGGAGAGGCTGTCCTGCTTCTGAGCTGCTGATCTCAT
CGGTCTTATAAACACTGACTCTTGGGATCATGGATCTGGTGTGATGGGACAGA
TTGGGAGAAGATCGTATGAGTGAGGCGCAAAGTTGGTCATCATGCTAAGGCAAG
GCCTTGGGAGGCGAGTGCAAGACCAGACCTGCTCCTGTTGTATAAACTCTCAAGAG
TCCGTTTTGA
CYP81D11 (from Arabidopsis) was isolated from plant cultivated in liquid culture. The truncated version of CYP81D11 (CYP81D11tr in italics) was created by deleting the hydrophobic N-terminal region.

```
ATGTCATCAACAAAGACAAATAATGGAAAAACTATATACCTAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC
```
CYP71D15 construct PM2-2 (limonene-3-hydroxylase from peppermint) was provided by Prof. Rodney Croteau (Institute of Biological Chemistry, Washington State University, Pullman, US).³⁶⁷

ATGGCTCTGTTATTAGCAGTTTTTTTTGTCCGCGCTTTATAATCTCTGTAAGTAACCT
ACACCATATCCCCCTCTTAATCACAACATGCGGAAAACCGGAACCCCAAGGAAAGTT
CCGGGCGGCCCCCGGCGAAGCGCGCCTGATCGGGCAACCCTACCTCCTCCTGTTGGGG
AAGCTCCGCACACGCTTGCTGCGCAGTGCGGTGGGGTAGTGCCACCCGCGGACCCG
ACGTGCAGCTGGGTGAGGTCTTCCGTCGTCCTTTGTCGCTGGAGGCAGGAGCACAA
CTGGGAGATGAGGTAGATGAGGCAGCAGAGATGAGACATCGGG
ACGAGGATCATGTGGTACGACAACGAGGACATCATCTTCAGCCCCTACAGCGAGC
ACTGGCGCCAGATGCGCAAGATCTGCGCTTCCGAGCTCTCTCTCTCCCCCCAGCT
CCGCCTCTCCGCCCTCTCCGCCTCTCCCCTCACCTCCGCTCCTCCGCCCTCAAG
CTGCTCCCTCGGACGGCGGCGCCCTGAGATGAGAGCTGCTGGCCCAGGGGCGGCCGG
ACGTGAGGATAGAGACGCTGACGTGCTCCATCATCTGCAGGGCGGCGTTCGGGAGCG
TGATCAGGGAACTTCGGGTTGGCAAACGTTGAGGTTCCATTGGCGCAGCTTCTTTACC
ACTTCGACTGGAAGTTGGCGGAAGGAATGAAACCTTCTGATATGGACATGTCTGA
GGCGGAAGGCCTTACCGGAATACTAAAGAACAATCTTCTTCTTGCCCCACCC
TACGATCCTTTCATCATGA
APPENDIX B

B2 Gene sequence used in Chapter 4

**ATR1tr** (P450 reductase from Arabidopsis) was isolated from plant cultivated in liquid culture. The truncated version of ATR1 was created by deleting the hydrophobic 46 amino acids of the N-terminal region.

```
TGGAAGAAAAACGACGCGGATCGGAGCGGGGAGCTGAAGCCTTTGATGATCCCTA
AGTCTCTTATGCTAAGGAGAAGATGATGATTGAGATTTGGAATCCGGAAGAC
TAGAGTCTCTATCTTTCTGTTAGCAGACTGGAAGCTGAAGCCTTTGATGATCCCTA
GCATTATCGAAGAAAATCAAGCAGAGCTTACACGCTGGAAGCTTTGATGATCCCTA
AACTTTGCGATTTCTGTTAAGGAGCTTTGATGATCAGAGCATTGAGGATGATTTTA
TGGAAGCTGGTTCTGCAAGATGACAGATGATTGTTGCAAGGAGACTTTGATGATCCCTA
```
ATR2tr (P450 reductase from Arabidopsis) was isolated from plant cultivated in liquid culture. The truncated version of ATR2 was created by deleting the hydrophobic 75 amino acids of the N-terminal region.

```
GGTTCTGGGAATTCAAAACGTGAGCCTCTTTAAGCCTTTGGTTATTAAGCCTC
GTAGGAAGAACTGCTGAGGACGTGCTGCCAGGAAGTTACCCTTCGTTCAACA
AACTTGGTACTGCTGAGGAGTGTGTTGCTAAAGAGCTTTAGGAGAGAGTTAGGCA
AGGAGAAAGCACAGGAGGAGGCTCTGTGAGGATGTGTTGCTAGATGACAT
TAGGAAAACAGACAAATGAGCAGTTTTTAATGAGGTTGCCCAGGTTGATGAC
AGCTTTCTGAGCAAGCAGCAGCGAGATTCTACAATGGTTCACC
GGAGGGAGATGACAGAGGAGAATGGCTTAAGAACTTGAAGTATGGAGTGTTTGGAT
TAGGAAACAGACAAATGAGCAGTTTTTAATGAGGTTGCCCAGGTTGATGAC
AGCTTTCTGAGCAAGCAGCAGCGAGATTCTACAATGGTTCACC
```

APPENDIX B
A2  Gene sequence used in Chapter 5 and 6

**ATR2tr** sequence originated from the *E. coli* expression codon optimised ATR2 gene sequence synthesised by GeneArt.

```
GGTAGCGGTAATAGCAAAACGTGTGAAACCCGTGAAACCCTGTTATTAACCGGC
GTGAAGAAGAATAAGCAGCTGCTGAAATAAAGTGACATTTTTTCTGCAACC
GACCACCGACAGAGGTACTTTCGTGAAATGAGACAGATGTCAGTACT
TGATGAAAACCCGCTTCAAAATTGATCAGTCGATATGATGATGCTACCC
AATGACGGATATGTGATTTGAGCTATGGAATAGCTGTTGGAACAGGGTCAGAC
CTGTGACTGATGATGTTTTTATAGCATTAGCAGCAGCTTCGAAAATTGCAGAACCCGTATTCATGTTACCTGTGCACTGGTGTATGAAAAAATGCCGACCGGTCGTATTCATAAAGGTGTTTGTAGCACCTGGATGAAAAATGCAGTGCCGTATGAAAAAAGCGAAAATTGTAGCAGCGCACCGATTTTTGTTCGTCAGAGCAATTTTAAACTGCCGAGCGATAGCAAAGTGCCGATTATTATGATTGGTCCGGGTACAGGTCTGGCACCGTTTCGTGGTTTTCTGCAAGAACGTCTGGCACTGGTTGAAA
GCGGTGTTGAACTGGGTCCGAGCGTTCTGTTTTTTGGTTGTCGTAATCGTCGCATGGATTTCATCTATGAAGAAGAACTGCAGCGTTTTGTGGAAAGCGGTGCACTGGCCGAACTGAGCGTTGCATTTAGCCGTGAAGGTCCGACCAAAGAATATGTTCAGCACAAATGATGGACAAAGCCAGCGATATTTGGAATATGATTAGCCAGGGTGCCTATCTGTATGTTTGTGGTGATGCAAAAGGTATGGCACGTGATGTTCATCGTAGCCTGCATACCATTGCACAAGAACAGGGTAGCATGGATAGCACCAAAGCCGAAGGTTTTGTATAAAAATCTGCAGACCAGCGGTCGTTATCTGCGTGATGTTTGGTAA
```
Modified P450 gene sequences for P450-ATR2tr fusion platform:

**IFS native** gene derived from Prof. Mattheos Koffas. The first six N-terminal residues (small letters) represent the synthetic mammalian peptide (amino acid sequence: MALLLAVF). The codon of the last amino acid (Serine) was changed from TCT to TCA to make it compatible for ligation independent cloning.

```
ATGCGGCTCTGTTATTAGCAGGTTTTCTTTCTGCACTTTGCGTCCACCAAGTGCAAAATCAAAGACTCTGCCCACCCTCCAAACCCCTCCAAGCCCAAGCTGCTTCTCTCTCTGACCTCTCAAGCGGACATGAGTCAAAATTACCAAGGAGCAATCAAGGGCCTTGTTGTCGACTTTTTCTCTGCAGGGACAGATTCCACAGCGGTGCACAACAGAGTGGGCATTGGCAGAGCTCATCAACAATCCCAGGGTGTTGCAAAAGGCTCGTGAGGAGGTCTACAGTGTTGTGGGCAAAGATAGACTCGTTGACGAAGTTGCACTCAAAACCTTCCTTACATTAGGGACATTGTGAAGGAGACATTCCGAATGCACCCACCACTCCCAGTGGTCAAAAGAAAGTGCACAGAAGAGTGTGAGATTAATGGGTAATGTGATCCCAGAGGGAGCATTGGTTCTTTTCAATGTTTGGCAAGTAGGAAGGGCCCCAAATACTGGGACAGACCATCAGAATTCCGTCCCGAGAGGTTCTTAGAAACTGGTGCTGAAGGGGAAGCAGGGCCTCTTGATCTTAGGGGCCAGCATTTCCAACTCCCTCCATTTGGGTCTGGGAGGAGAATGTGCCCTGGTGTCAATTTGGCTACTTCAGGAAATGGCAACACTTCTTGCATCTCTTATCCAATGCTTTGACCTGCAAGTGCTGGGCCTCAAGGACAAATATTGAAAGGTGATGATGCCAAAGTTTAGCATGGAAGAGAGACCTGGCCTCACAGTTCCAAGGGCACATAGTCTCGTTTGTGTTCCACTTGCAAGGATCGGCGTTGCATCTAAACTCCTTCTCA
```
**IFSopt** gene was synthesised and codon optimised for the expression in *E. coli* by GeneArt. The codon of the last amino acid (Serine) was changed from AGC to TCA to make it compatible for ligation independent cloning.

```plaintext
GCCGTCGCCAGCCGAGCGCAAAAAAGCAAGCAGCTGCGTATCTGCCGAATCCGCCTAGCCTGGTTATAGGCAGCAGCAAAACATGTTGCGTGGTTAATCCGCTTCTGCATATGAGCAAGCAAAGCAGACAGTGGTCCGCTGTTAGCTGAGCTTTGGTAGCATGCCGACCGTTGTTGCAAGCACACCGGAACTGTTTAAAC
TGTTTTCGCTGACCGATCTGCAAGCAGCGCACAAGCTGTTGGAATTTTTTGCTGGTTTTTCTGGACACCATGGAAATTAAAATTAACCAAGAAACATTAAAGGCCCTGTTGCTGTTAGATATTTGCTGAAATATAGCGTTGCAATGGTTCCGGTGTTGCTTGATGAGCAGAAAGCACAGAAACCGCTGGATGTTACCGAGGAACTGCTGAAATGGACCAATAGCACCATTA
GCATGATGATGCGGTAAGGCAGCAGACTGAGCAGCTCATCTGCCGCTGCGTACCCAGCAGATTCGTAAATTTCTGCGTGTTATGGCAGAGCGCAGAAGCACAGAAACCGCTGGATGTTACCGAGGAACTGCTGAAATGGACCAATAGCACCATTA
GCATGATGATGCTGGGTGAAGCCGAAGAAATTCGTGATA
```
CYP73A5 (cinnamate-4-hydroxylase from Arabidopsis) was synthesised and codon optimised for the expression in *E. coli* by GeneArt. The truncated version of CYP73A5 (73A5tr) was created by deleting the hydrophobic N-terminal region and to improve the expression in *E. coli* and Alanin was added just after the start codon methionin (in bold).\textsuperscript{321}

The codon of the last amino acid (Cysteine) will be changed from TGC to TCA (Serine) to make it compatible for ligation independent cloning.
**APPENDIX B**

**CYP82E4tr** (N-demethylase from *Nicotiana tabacum*) derived from Prof. Ralph Dewey. The truncated version of CYP82E4 (82E4tr) was created by deleting the hydrophobic N-terminal region. The codon TAT of the last amino acid (Tyrosine) will be changed from to ACA (Threonine) to keep the electrophilic charge and to make it compatible for ligation independent cloning.

```
ACAAAAAAATCTCAAAAACCTTACCCACCCGAAAATCCCGGAGGAT
GGCCGGTAATCGGCCATTTTCCACTTCAATGACGCGACGCCGACCTTCTATT
AGCTCGAAACTCGGAGACTTAGCTGACAAATACGGCCCCGTTTTTCTCTTGCG
CTAGGGCTCTCCCTGGTCTTAGTTGTAAGCGTTAAGCTGAAACTGCTGGT
TCTCTACAAATGACGCCATTTTCTGAGTCAGCTGAGCCGATTACCTTGCG
AAATGGACAAATTTTAGTTATTCAGGAAGTTCTCTGCTAGCTGCTGCTCG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
AAAAATTCGAAAAACGGCTATTTTCTGAAAGATATGATAGATAGATAGATAG
CTGAGGAAGATATTTTCTGTAATGACGACCTTCTGTAATTTTCTGTAATGACG
AGGCTTCTGATCTGAGATACGTGTCTGAGGCTTCTGCTGAGGCTTCTGCTG
```
CYP81D8 (from Arabidopsis) was isolated from plant cultivated in liquid culture. The truncated version of CYP81D8 (81D8tr) was created by deleting the hydrophobic N-terminal region. The codon GTT (Valine) was changed to GTA (Valine) to make it compatible for ligation independent cloning.

GGAAAACCTCAAGCGAAGCCAAATCTACCTCCGAGTCCGGCATGGTCATTACCGG
TGATTTGTCATCTCCGCTTTTCTCAAAACCACCGATTCAGCACATCTCTTCTCTCTCT
CTCTCAATCTTAAACATGCTCCGATCTTCCTCCGCTACGTGTAACCCGACTC
GGTTTCTGAACTCTGTCACACTCGAGTCCGCCAGGAATGTGTTCACCAAAAACCGACG
TCGTACCTGCGCAACAGACCATAACACCTACCTCCGGATTA
CACAACCATGATCAGCTTCTACGTTGACACTGCGTAACCTCCGCGCATC
GGCTCCGTGAGATATTTCTCCAAATCACGCTCAATAGCCTTTTGTCTATTCGTA
AAGACGAGATCCGACGACTTGTGGTTTCGTCTTTCTCGAAACTTTTTCACAAGAGTT
TGTTGAAGTGGATATGAAATCAATGTTATCTGACTTAACATTCAACAACATTTTTA
AGAATGTGTCGCGGAAAACGTTAATCAGCGAGACTGGAGATGTTGAGATCGAGCGGAGG
CTAAACGGTCTCCGCGCAGCTTTATAGCGGATGTGGTGTGGGTGCTGCTGAAAC
CGCTGGATTCAATTACCGGTTTTTTCGGTGTTTCGATACGAGACACGGGGTT
AAGAAGTTGCGGGTAGGCTCGAGCTGGTTTCTCGAAGGTGTTGTTGAGAAAC
GAGAAGCTAAGGAGAAAGAAACACTATGATCGCTACCTTGTTCTCGCAAGAG
ATCCCAACCGGATATTACCTACGAGATCTCATATTAAGGAAAACATGCTCGCTTTTG
ATACCTGACAGGGACAGCAACATACGGGTTACGGTAGAGATGGCATTGCGGAGG
TGTTGAACCACATCCCGATGTATTTGAAACAGGGCAGAGATGAAACTCGATAGAAAGAT
AGGTTTAGACAGCGCTTATGGAATACGATATCTCAAAACCTTTATCTCCAA
AACATTGTTCTGAAACGTTTGCGCCCTTTATCTTGCCGCTACGTCTTCTTCTC
ACGTTTGCTCGGAAGATGGTAGAAGTTGGCACGATACGATATCGCGGTCGAGCAGAT
ACTATGACCAATGTTGGCGCTATACACAGAGATCTACGATAGATGCCATCCCA
ATGAGCTTCAAGCCAGAGAGGTGTGAGAAGAAGAAGAGAGCAGCTCAGAAGCTATGC
CGTTTGGTTAGGAAGAAGGGGTGTGCTTGTCTGACTGTGCTTCGGCTCTTAT
AAACCTCGACTTGGAGATATGCTGTGGTTGGAATGGGAGAGATTGGAGAA
GAAGTGGATATGAGTTGAAGGGCAAGTGTTTCATGGCAATGCTATTAAAGGCGAGGTTGG
AAGCCATGTCGAGACGACGTCCTCT GTTA AAAATCTTCAACGAGTCCGTA
Abbreviations

½ MS Murashige and Skoog medium half strength
2-ADNT 2-amino-4,6-dinitrotoluene
2-HADNT 2-hydroxylamino-4,6-dinitrotoluene
4-ADNT 4-amino-2,6-dinitrotoluene
4-HADNT 4-hydroxylamino-2,6-dinitrotoluene
A adenine
aa amino acids
Abs optical absorbance
ADE 2d selection marker for adenine auxotrophy
ALA 5-aminolevulinic acid
ampR ampinillin resistance gene
ATR1 Arabidopsis thaliana cytochrome P450 reductase 1
BLAST Basic Local Alignment Search Tool
bp base pairs
BSA bovine serum albumine
C cytosine
cDNA complementary DNA
CO carbon monoxide
CPO chloroperoxidase
CPR cytochrome P450 reductase
DMSO dimethylsulfoxide
DNA deoxyribonucleic acid
dNTP dinucleotide triphosphate
DTT dithiothreitol
[E] enzyme concentration
EDTA ethylenediaminetetraacetic acid
ER endoplasmic reticulum
et al. et alii
EtOH ethanol
f1 ori f1 phage origin of replication
FAD flavin adenine dinucleotide
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FMN</td>
<td>flavin mononucleotide</td>
</tr>
<tr>
<td>g</td>
<td>grams</td>
</tr>
<tr>
<td>G</td>
<td>guanine</td>
</tr>
<tr>
<td>heme</td>
<td>Protoporphin IX bound to one iron atom</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazinethanesulphonic acid</td>
</tr>
<tr>
<td>His</td>
<td>6x histidines residues</td>
</tr>
<tr>
<td>HiTEL</td>
<td>High Throughput Expression Laboratory</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>K$_2$HPO$_4$</td>
<td>dipotassium hydrogen phosphate</td>
</tr>
<tr>
<td>kanR</td>
<td>kanamycine resistance gene</td>
</tr>
<tr>
<td>kb</td>
<td>kilo base pairs</td>
</tr>
<tr>
<td>$k_{cat}$</td>
<td>maximum turnover rate</td>
</tr>
<tr>
<td>KCl</td>
<td>potassium chloride</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodaltons</td>
</tr>
<tr>
<td>kg</td>
<td>kilogram</td>
</tr>
<tr>
<td>KH$_2$PO$_4$</td>
<td>potassium dihydrogen phosphate</td>
</tr>
<tr>
<td>$K_M$</td>
<td>Michaelis-Menten constant</td>
</tr>
<tr>
<td>l</td>
<td>litre</td>
</tr>
<tr>
<td>LacI</td>
<td>repressor gene for IPTG induction</td>
</tr>
<tr>
<td>LacZ</td>
<td>β-galactosidase gene</td>
</tr>
<tr>
<td>LB</td>
<td>Lysogeny Broth</td>
</tr>
<tr>
<td>LIC</td>
<td>Ligation independent cloning</td>
</tr>
<tr>
<td>LiCl</td>
<td>lithium chloride</td>
</tr>
<tr>
<td>LIC-vector</td>
<td>pETYSBLIC3C vector</td>
</tr>
<tr>
<td>m</td>
<td>meter</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>M13 ori</td>
<td>origin of replication of filamentour phage M13</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>Matrix-Assisted Laser Desorption/Ionization - <em>Time-of-Flight</em></td>
</tr>
<tr>
<td>MgSO$_4$</td>
<td>magnesium phosphate</td>
</tr>
<tr>
<td>min</td>
<td>minute(s)</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
</tbody>
</table>
ABBREVIATIONS

NAD\(^+\) nicotinamide adenine dinucleotide, oxidised form
NADH nicotinamide adenine dinucleotide, reduced form
NADP\(^+\) nicotinamide adenine dinucleotide phosphate, oxidised form
NADPH nicotinamide adenine dinucleotide phosphate, reduced form
NaOH sodium hydroxide
NCBI National Center for Biotechnology Information
NDSB 201 3-(1-Pyridinio)-1-propanesulfonate
NEB New England Biolabs
(NH\(_4\))^2SO\(_4\) ammonium sulphate
NOS nitric oxide synthase
OD\(_{600}\) optical density at 600 nm
oLAC Lac operator
PAGE polyacrylamide gel electrophoresis
pBR322 ori origin of replication of the plasmid pBR322
PBS phosphate buffered saline
PCR polymerase chain reaction
PEG poly-ethylene glycol
pGAL galactose promoter
PMSF phenylmethanesulfonylfluoride
pT7 T7 promotor
RBS ribosome binding site
RDX hexahydro -1,3,5-trinitro -1,3,5-triazine, Royal Demolition Explosive
RNA ribose nucleic acid
rpm revolutions per minute
RT reverse transcriptase
[S] substrate concentration
SDS sodium dodecyl sulfate
sec second(s)
SGI synthetic media for yeast
SRS substrate recognition site
SSNMR Solid State Nuclear Magnetic Resonance
STDEV one standard deviation from the mean
T Thymine
ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>tac</td>
<td>tac promoter</td>
</tr>
<tr>
<td>TCA</td>
<td>trichloroacetic acid</td>
</tr>
<tr>
<td>TE</td>
<td>Tris/EDTA</td>
</tr>
<tr>
<td>Temed</td>
<td>N, N, N’, N’-Tetramethylethylendiamine</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
<tr>
<td>TNT</td>
<td>2,4,6-trinitrotoluene</td>
</tr>
<tr>
<td>tPGK</td>
<td>phosphoglycerate kinase terminator</td>
</tr>
<tr>
<td>Tris</td>
<td>2-amino-2-hydroxymethyl-1,3-propanediol</td>
</tr>
<tr>
<td>Triton X-100</td>
<td>4-octylphenol polyethoxylate</td>
</tr>
<tr>
<td>tT7</td>
<td>T7 terminator</td>
</tr>
<tr>
<td>URA3</td>
<td>selection marker for uracil auxotrophy</td>
</tr>
<tr>
<td>UV</td>
<td>ultra violet (light)</td>
</tr>
<tr>
<td>v/v</td>
<td>volume to volume ratio</td>
</tr>
<tr>
<td>V_max</td>
<td>maximal reaction velocity</td>
</tr>
<tr>
<td>w/v</td>
<td>weight to volume ratio</td>
</tr>
<tr>
<td>X-Gal</td>
<td>bromochloroindolylgalactopyranoside</td>
</tr>
<tr>
<td>YPGA</td>
<td>yeast medium containing yeast extract, bactopeptone, glucose and adenine</td>
</tr>
<tr>
<td>YPGE</td>
<td>yeast medium containing yeast extract, bactopeptone, glucose and ethanol</td>
</tr>
<tr>
<td>ZnCl₂</td>
<td>zinc chloride</td>
</tr>
</tbody>
</table>
References


REFERENCES


61 Nielsen, K. A. & Möller, B. L. in Cytochromes P450: Structure, Mechanism, and Biochemistry (ed P. R. Ortiz de Montellano) Ch. 12, 553-583 (Kluwer Academic/Plenum Publisher, 2005).


Sauveplane, V. et al. Arabidopsis thaliana CYP77A4 is the first cytochrome P450 able to catalyze the epoxidation of free fatty acids in plants. *FEBS J* **276**, 719-735 (2009).


References


253 Kim, G. T. et al. CYP90C1 and CYP90D1 are involved in different steps in the brassinosteroid biosynthesis pathway in Arabidopsis thaliana. Plant J 41, 710-721 (2005).


REFERENCES


Schalk, M. & Croteau, R. A single amino acid substitution (F363I) converts the regiochemistry of the spearmint (-)-limonene hydroxylase from a C6- to a C3-hydroxylase. Proc Natl Acad Sci U S A 97, 11948-11953 (2000).


REFERENCES


245


