

Characterisation of Novel Cytochrome P450 Fusion Systems

A thesis submitted to the University of Manchester for the
degree of Doctor of Philosophy in the Faculty of Life Sciences.

2010

Jacob W. Robinson

Front Matter

List of Contents

Front Matter	3
List of Contents	3
List of Figures	8
List of Tables	15
List of Abbreviations	16
Abstract	19
Declaration	20
Copyright Statement	20
Acknowledgements	21
1. Introduction	25
1.1. Literature review	25
1.1.1. What are enzymes?.....	25
1.1.2. Roles of cofactors in enzyme reactions.....	25
1.1.3. The Cytochromes P450.....	32
1.1.4. P450 Nomenclature.....	32
1.1.5. Relevance of P450s to Human Health	33
1.1.6. The reactions catalysed by P450s	35
1.1.7. Structure of P450s	39
1.1.8. Redox partners	41
1.1.9. The History of Fusion P450s	42
1.1.10. Engineering BM3	46
1.2. Discussion of theory behind the techniques used	48
1.2.1. Enzyme Kinetic Studies.....	48
1.2.2. Ligand Binding.....	50
1.2.3. Resonance Raman Spectroscopy	52
1.2.4. Electron Paramagnetic Resonance	54
1.2.5. Redox Potentiometry.....	56
1.2.6. Differential Scanning Calorimetry	58
1.3. Enzyme Systems Studied	59

1.3.1.	CYP102A3	59
1.3.2.	CYP116B1	61
1.4.	Experimental Aims and Thesis Outline	62
2.	Materials and Methods	67
2.1.	General Information	67
2.2.	Materials used	67
2.2.1.	Microbiology Reagents	67
2.2.2.	Buffers	68
2.2.2.1.	Tris EDTA (TE) buffers	68
2.2.2.2.	Phosphate Buffers	68
2.2.2.3.	Nickel Purification Buffers	68
2.2.2.4.	Assay Buffers	68
2.2.3.	Protein Chromatography	69
2.2.3.1.	Columns	69
2.2.3.2.	Linear gradients	70
2.3.	Suppliers	71
2.3.1.	CYP116B1 Substrate Candidates.....	71
2.3.2.	Cell Strains.....	71
2.3.3.	Enzymes	72
2.4.	Experimental Methods	72
2.4.1.	Site-Directed Mutagenesis	72
2.4.2.	Ligation Reactions.....	72
2.4.3.	Transformation of Plasmid DNA into <i>E. coli</i> Cells.....	73
2.4.4.	Glycerol Stocks	74
2.4.5.	Amplification of Plasmid DNA	74
2.4.6.	Protein Expression Trials	75
2.4.7.	Expression and Purification of HDCYP102A3	75
2.4.8.	Expression and Purification of FLCYP116B1 and HDCYP116B1.....	77
2.4.9.	Expression and Purification of RDCYP116B1.....	78
2.4.10.	Multi-Angle Light Scattering (MALS).....	79
2.4.11.	Crystallisation Trials.....	80
2.4.12.	Binding Titrations	81
2.4.13.	Redox Potentiometry.....	81
2.4.14.	Reduced-CO Complex Formation.....	82
2.4.15.	Preparation of FLCYP116B1 for kinetic assays	83
2.4.16.	Steady State Assays	83
2.4.17.	Electron Paramagnetic Resonance Spectroscopy.....	85
2.4.17.1.	Preparation of EPR Samples	85
2.4.17.2.	EPR of HDCYP102A3.....	86

2.4.17.3.	EPR of CYP116B1 and its Haem Domain.....	86
2.4.18.	Differential Scanning Calorimetry.....	87
2.4.19.	Resonance Raman Spectroscopy	88
2.4.20.	Product Determination from Prospective Substrate Oxidation by CYP116B1	88
2.4.20.1.	CYP116B1 Reaction	88
2.4.20.2.	Product Purification	89
2.4.20.3.	High Performance Liquid Chromatography-Mass Spectrometry.....	89
3.	<i>The Haem Domain of CYP102A3</i>	93
3.1.	General Information	93
3.2.	Results.....	98
3.2.1.	Expression and Purification of CYP102A3 haem domain.....	98
3.2.2.	Ligand binding studies.....	100
3.2.2.1.	Substrates.....	101
3.2.2.2.	Inhibitors.....	108
3.2.3.	Electron Paramagnetic Resonance (EPR) spectroscopy.....	113
3.2.4.	Resonance Raman spectroscopy	115
3.2.5.	Redox potentiometry.....	119
3.2.6.	Differential Scanning Calorimetry	124
3.2.7.	Chemical denaturation.....	126
3.2.8.	Light Scattering.....	130
3.2.9.	Crystallographic trials.....	132
3.2.10.	Homology modelling.....	134
3.3.	Summary & Discussion	137
4.	<i>CYP116B1 Protein Studies.....</i>	143
4.1.	General information	143
4.1.1.	Survey of P450 Fusions.....	143
4.1.1.1.	P450-Redox Partner Fusions.....	143
4.1.1.2.	P450-Known Enzyme Fusions.....	146
4.1.1.3.	P450-Other Domain Fusions	149
4.1.2.	CYP116B1	150
4.2.	Cloning	152
4.2.1.	FLCYP116B1	152
4.2.2.	HDCYP116B1.....	160
4.2.3.	RDCYP116B1.....	161
4.3.	Expression and purification.....	163
4.3.1.	FLCYP116B1 and HDCYP116B1	163
4.3.2.	RDCYP116B1	166

4.4.	Binding of Small Ligands	168
4.5.	Electron Paramagnetic Resonance Spectroscopy.....	171
4.5.1.	FLCYP116B1.....	171
4.5.2.	HDCYP116B1	175
4.6.	Differential Scanning Calorimetry	177
4.6.1.	HDCYP116B1	177
4.6.2.	FLCYP116B1.....	179
4.7.	Multi-Angle Light scattering.....	181
4.7.1.	FLCYP116B1.....	181
4.7.2.	HDCYP116B1	183
4.8.	Crystallographic trials.....	186
4.9.	Homology Modelling	186
4.9.1.	HDCYP116B1	186
4.9.2.	RDCYP116B1.....	188
4.10.	Summary & Discussion.....	191
5.	<i>CYP116B1 Enzyme Studies</i>	197
5.1.	General Information	197
5.1.1.	Bioremediation with P450s.....	198
5.1.2.	Experimental Outline.....	199
5.1.3.	Substrate Library.....	199
5.2.	Binding Studies	203
5.2.1.	Substrate Binding.....	203
5.2.2.	Inhibitor Binding.....	204
5.3.	Redox Potentiometry.....	207
5.4.	Steady State Kinetics of FLCYP116B1	210
5.5.	Product Determination.....	215
5.5.1.	Atrazine	216
5.5.2.	Di-allate.....	218
5.5.3.	EPTC.....	221
5.5.4.	Molinate.....	223
5.5.5.	Prosulfocarb	225
5.5.6.	Thiobencarb.....	228
5.5.7.	Vanillic acid.....	231
5.5.8.	Vernolate.....	233
5.6.	Steady State Kinetics of RDCYP116B1	235

5.7. Summary & Discussion	242
6. Summary and Bibliography.....	249
6.1. Summary	249
6.1.1. HDCYP102A3.....	249
6.1.2. CYP116B1	250
6.1.3. Synthetic Chemistry with P450s.....	251
6.1.4. Bioremediation Using P450s.....	251
6.1.5. Suggestions for Future Studies.....	252
6.2. Bibliography	255

Word count: 62555.

List of Figures

Figure 1.1. Structure of acetyl coenzyme A (CoA).....	26
Figure 1.2. Full structure of oxidised nicotinamide adenine dinucleotide phosphate (NADP ⁺).	27
Figure 1.3. Structure of the flavin adenine dinucleotide (FAD) cofactor.	28
Figure 1.4. A selection of biologically occurring iron-sulfur clusters.....	29
Figure 1.5. Structures of two [2Fe-2S] iron sulfur clusters with the side chains of ligating amino acid residues and inorganic sulfurs between iron atoms.	30
Figure 1.6. Structures of haems A, B and C.....	31
Figure 1.7. Structural representations of the P450 haem environment	32
Figure 1.8. The classical P450 reaction.....	35
Figure 1.9. Diagram representing the d-electron configuration of ferric iron (Fe ³⁺) ...	36
Figure 1.10. Schematic diagram of the catalytic cycle of a cytochrome P450.	38
Figure 1.11. Proposed radical rebound mechanism describing the terminal oxidation event in the P450 catalytic cycle.....	39
Figure 1.12. Logo showing the consensus amino acid sequence of the P450 haem binding motif.....	40
Figure 1.13 Structures of cytochrome P450 enzymes from various organisms showing structural similarity.	41
Figure 1.14. Phylogenetic tree showing all presently sequenced members of the CYP102A subfamily.	45
Figure 1.15. Schematic representations of currently confirmed P450 fusions.....	46
Figure 1.16. Diagram showing energetic transitions involved in Rayleigh and Raman scattering models.	53
Figure 1.17. Representative diagram showing the splitting, due to the Zeeman effect, of the spin moments of electrons in an applied magnetic field B_0	55
Figure 1.18 Image of phthalate dioxygenase reductase (PDOR) from the crystal structure.....	62
Figure 2.1. Schematic diagram of a gradient mixer.	71
Figure 2.2. Representative diagram of a well unit from a 96-well crystallography tray	80
Figure 2.3. Representative diagram of a basic stopped flow apparatus.....	85

Figure 2.4. Schematic, labelled diagram of a differential scanning calorimeter	87
Figure 3.1. Section of an amino acid sequence alignment of members of CYP102A P450 subfamily	95
Figure 3.2. Sections of an amino acid sequence alignment of members of CYP102A P450 subfamily	96
Figure 3.3. Image of an SDS-PAGE gel (12%) showing samples of the haem domain of CYP102A3	98
Figure 3.4. UV/Visible spectra showing the changes in absorbance upon reduction and binding of CO to the iron of the haem domain of CYP102A3.....	100
Figure 3.5. UV/visible spectra of the titration of 13-methylmyristic acid.....	102
Figure 3.6. Difference spectra for the titration of 13-methylmyristic acid	102
Figure 3.7. Binding curves for HDCYP102A3 (~5 μ M) against a selection of fatty acid substrates.....	104
Figure 3.8. Binding curves for HDCYP102A3 (~5 μ M) against three fatty acid substrates:	105
Figure 3.9. Structures ofazole inhibitors used for binding titration experiments.	108
Figure 3.10. UV/visible spectra of the titration of 4-phenylimidazole.....	109
Figure 3.11. Difference spectra for the titration of 4-phenylimidazole.....	110
Figure 3.12. Binding curves for a selection of inhibitors.....	111
Figure 3.13. EPR spectrum of the HD CYP102A3 (200 μ M) in the oxidised state recorded at 10K.....	114
Figure 3.14. Resonance Raman spectra of HDCYP102A3.	116
Figure 3.15. Sections (1750 – 1325 cm^{-1}) of resonance Raman spectra	116
Figure 3.16. Sections (900 – 200 cm^{-1}) of resonance Raman spectra.....	117
Figure 3.17. UV/Visible spectra of HDCYP102A3 (~6 μ M) under anaerobic conditions during reduction with sodium dithionite.	120
Figure 3.18. UV/Visible spectra of HDCYP102A3 under anaerobic conditions during sequential reduction with sodium dithionite in the presence of substrate <i>N</i> -palmitoyl glycine (NPG).....	121
Figure 3.19. Difference spectra showing the wavelengths of maximum change during sequential reduction of HDCYP102A3.....	122
Figure 3.20. Redox potential curves for the haem domain of CYP102A3	123
Figure 3.21. DSC trace for HDCYP102A3.....	126

Figure 3.22. UV/Visible Spectra showing changes in absorbance of HDCYP102A3 (~1.8 μ M) following incubations with different concentrations of the denaturant guanidinium chloride.	127
Figure 3.23. Difference spectra showing the change in absorbance of HDCYP102A3 (~1.8 μ M) on incubation with varying concentrations of the chaotropic denaturing agent guanidinium chloride (GdmCl).	128
Figure 3.24. Denaturation curve for HDCYP102A3.....	129
Figure 3.25. Trace from a Multi angle light scattering (MALS) run of a sample of HDCYP102A3.....	131
Figure 3.26. A plot of values, derived from Zimm fitting, of the hydrodynamic radius (R_h) and molecular mass against elution volume for a multi angle light scattering (MALS) run of a sample of HDCYP102A3.	132
Figure 3.27. Images of the HDCYP102A3 crystals obtained.....	133
Figure 3.28. Side by side comparison of the template structure of BM3 haem domain (blue) and the model of HDCYP102A3.....	134
Figure 3.29. Close-up image of the homology model of HDCYP102A3 (green) with the template structure (HD-BM3) overlaid	135
Figure 3.30 Close-up image of the homology model of HDCYP102A3 (green) with the template structure (HD-BM3) overlaid	136
Figure 4.1. Phylogenetic tree of amino acid sequences listed as P450-phthalate dioxygenase reductase-like fusions.....	145
Figure 4.2. Domain schematic for the P450-acylCoA dehydrogenase (ACAD) fusion (CYP221A1) from <i>Pseudomonas fluorescens</i>	148
Figure 4.3. Plasmid map of a pGEM-T construct containing a PCR fragment of the gene encoding CYP116B1	153
Figure 4.4. Plasmid map of the vector pET-15b supplied by NEB.	154
Figure 4.5. Images of 0.8 % agarose gels showing pET-15b vector	156
Figure 4.6. Diagram showing the recognition sites and cutting patterns for AseI and NdeI restriction enzymes.....	157
Figure 4.7. a) Plasmid map of the pET-15b construct containing the insert of the <i>CYP116B1</i> gene from <i>C. metallidurans</i>	157
Figure 4.8. Agarose gel (0.8%) showing miniprep DNA (NcoI undigested and digested) from ligation reactions	159
Figure 4.9. DNA Sequencing chromatogram showing the pET-15b start codon	159

Figure 4.10. Beginning of sequencing chromatogram for the pHDCYP116B1-15b candidate.....	161
Figure 4.11. Section of an alignment of the CYP116B1 gene sequence (1 360 – 1413 b.p.) and the resulting sequence from DNA sequencing of the pRDCYP116B1-11a construct.....	161
Figure 4.12. Side by side comparison of the sequences of the protein that would be expressed from the pRDCYP116B1-11a construct.....	162
Figure 4.13. Section of the sequencing chromatogram showing the T7 forward sequence of the pRDCYP116B1-11a construct.....	163
Figure 4.14. Images of SDS PAGE gels showing samples of expression trial cultures at time intervals with and without IPTG induction.	164
Figure 4.15 Image of an SDS-PAGE gel showing the purification of FLCYP116B1 using a 6xHis tag affinity column.....	165
Figure 4.16 The UV/visible spectrum of purified RDCYP16B1 protein.....	167
Figure 4.17. UV/Visible spectra of FLYCYP116B1	168
Figure 4.18. Difference spectra (with original spectra as an inset) showing the binding titration of cyanide, in the form of NaCN (200 – 2000 μ M), against HDCYP116B1 (\sim 3 μ M).....	169
Figure 4.19. Binding curve for the titration of NaCN against HDCYP116B1 (\sim 3 μ M).	170
Figure 4.20. Electron paramagnetic resonance spectra of FLCYP116B1.....	172
Figure 4.21. Expansion of the EPR spectrum (2500 – 3600 Gauss) of oxidised FLCYP116B1	173
Figure 4.22. Expansion of the EPR spectrum of dithionite reduced FLCYP116B1	174
Figure 4.23. EPR spectrum of oxidised HDCYP116B1 (200 μ M) purified without imidazole.....	176
Figure 4.24. EPR spectrum of oxidised HDCYP116B1 (200 μ M) purified without imidazole, and with 1 mM imidazole added.	177
Figure 4.25. Plot of results of a DSC experiment on HDCYP116B1 (\sim 20 μ M).	178
Figure 4.26. Plot of the results of a DSC experiment on HDCYP116B1 (\sim 20 μ M) in the presence of 2-phenylimidazole (1 mM).....	178
Figure 4.27. Plot of the results of a DSC experiment on FLCYP116B1 (\sim 20 μ M).....	180
Figure 4.28. Plot of the results of a DSC experiment on FLCYP116B1 (\sim 20 μ M) in the presence of 2-phenylimidazole (1 mM).....	180

Figure 4.29. Trace from a MALS run of FLCYP116B1.....	182
Figure 4.30. Plots of values of hydrodynamic radius (R_h) and molecular mass for each of the three peaks in the MALS trace for FLCYP116B1.....	183
Figure 4.31. Trace from a MALS run of HDCYP116B1	184
Figure 4.32. Plots of values of hydrodynamic radius (R_h) and molecular mass for the peak in the MALS trace of HDCYP116B1.....	185
Figure 4.33 Image of a homology model of the haem domain of CYP116B1, shown as a cartoon.....	187
Figure 4.34 Image of a homology model of the haem domain of CYP116B1, shown as a cartoon.....	188
Figure 4.35 Image of a homology model of RDCYP116B1 shown as a cartoon.....	189
Figure 4.36. Images of the homology model of the reductase domain of CYP116B1, focussing on the redox cofactors.....	190
Figure 5.1. Proposed mechanism of the degradation of EPTC involving CYP116 enzymes.....	198
Figure 5.2. Structures of the eight CYP116B1 substrate candidates	202
Figure 5.3. UV/visible spectra showing slight changes in haem Soret band of HDCYP116B1 (4 - 6 μ M) in the presence of substrates.....	203
Figure 5.4. Binding curve for imidazole (5 - 1950 μ M) titrated against HDCYP116B1	204
Figure 5.5. Binding data for imidazole titrated against HDCYP116B1	205
Figure 5.6. Binding curves for HDCYP116B1 (~5 μ M) with (left) 2- and (right) 4-phenylimidazoles.....	206
Figure 5.7. Redox potential curves for the haem domain of CYP116B1.....	209
Figure 5.8. Kinetic curves for the dependency of FLCYP116B1-dependent NADPH oxidation on the substrate concentration	211
Figure 5.9. HPLC traces for the products of the reaction of atrazine with CYP116B1.....	216
Figure 5.10. Mass spectrum of the purified products of a CYP116B1 turnover reaction in the presence of substrate candidate atrazine.....	217
Figure 5.11. HPLC traces for the products of the reaction of di-allate with CYP116B1.....	218
Figure 5.12. Mass spectrum of the purified products of a CYP116B1 turnover reaction in the presence of substrate candidate di-allate.....	219

Figure 5.13 Mass spectrum of the purified products of a CYP116B1 turnover reaction in the presence of substrate candidate di-allate.....	220
Figure 5.14. HPLC traces for the products of the reaction of EPTC with CYP116B1.	221
Figure 5.15. Mass spectrum of a CYP116B1 turnover reaction in the presence of substrate EPTC.	222
Figure 5.16 HPLC traces for the products of the reaction of molinate with CYP116B1.	223
Figure 5.17. Mass spectrum of a CYP116B1 turnover reaction in the presence of substrate molinate.	224
Figure 5.18. HPLC traces for the products of the reaction of prosulfocarb with CYP116B1.	225
Figure 5.19. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate prosulfocarb.	226
Figure 5.20. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate prosulfocarb.	227
Figure 5.21. HPLC traces for the products of the reaction of thiobencarb with CYP116B1.	228
Figure 5.22. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate thiobencarb.....	229
Figure 5.23. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate thiobencarb.....	230
Figure 5.24. HPLC traces for the products of the reaction of vanillic acid with CYP116B1.	231
Figure 5.25. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate vanillic acid.	232
Figure 5.26. HPLC traces for the products of the reaction of atrazine with CYP116B1.	233
Figure 5.27. Mass spectrum of a CYP116B1 turnover reaction against substrate vernolate.....	234
Figure 5.28. Kinetic curves showing the dependency of the rate of NADPH oxidation on the concentration of electron acceptors.....	237
Figure 5.29. Images of the aligned structures of spinach ferredoxin reductase.....	241
Figure 5.30. Close-up view of the structures shown in Figure 5.29.	241

List of Tables

Table 1.1. List of enzymes in the CYP102A subfamily	44
Table 2.1. Table of antibiotics used	67
Table 2.2. Information about the chromatography columns.....	70
Table 2.3. List of mediator compounds.....	82
Table 3.1. List of known enzymes of the CYP102 family	94
Table 3.2. Dissociation constants for HDCYP102A3.....	107
Table 3.3. Dissociation constants for HDCYP102A3 with several inhibitor ligands ..	113
Table 3.4. g-values from electron paramagnetic resonance spectroscopy	115
Table 3.5. Selected vibrational bands	118
Table 4.1. Selected results from a CDART search	144
Table 4.2. Selected results from a CDART search	147
Table 4.3. Selected results from a CDART search	149
Table 4.4. g-values from electron paramagnetic resonance spectroscopy	174
Table 5.1. Library of potential substrates for CYP116B1.....	201
Table 5.2. Dissociation constants (K_d/K_H)	207
Table 5.3. Kinetic constants (k_{cat} and K_M) for FLCYP116B1	213
Table 5.4. Kinetic constants (k_{cat} and K_M) for the RDCYP116B1 reaction.....	238
Table 5.5. Kinetic constants (k_{cat} and K_M) for the RDCYP116B1 reaction.....	240
Table 5.6. Collated data for substrate candidates	244

List of Abbreviations

Abs	Absorbance
ACAD	Acetyl coenzyme A dehydrogenase
ALDH	Aldehyde dehydrogenase
AMP	Adenosine monophosphate
b.p.	Base pairs
BDOR	Benzoate dioxygenase reductase
BM3	P450 BM3 (CYP102A1)
BV	Benzyl viologen
CAD	Cinnamyl alcohol dehydrogenase
CAH	Congenital adrenal hyperplasia
CDART	Conserved domain architecture retrieval tool
CFSE	Crystal field stabilisation energy
CoA	Coenzyme A
CPR	Cytochrome P450 reductase
CYP	Cytochrome P450
DCPIP	2,6-dichlorophenolindophenol
DEAE	Diethyl aminoethyl cellulose
dH ₂ O	distilled deionised water
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylene diamine tetraacetic acid
EPR	Electron paramagnetic resonance
EPTC	<i>S</i> -ethyl dipropylthiocarbamate
FAD	Flavin adenine dinucleotide
Fd	Ferredoxin
FdR	Ferredoxin reductase
FLCYPXXX	Full-length CYPXXX
FMN	Flavin mononucleotide
FNR	Ferredoxin reductase family reductase
GdmCl	Guanidine chloride
GSH	Glutathione
GST	Glutathione-S-transferase
HDCYPXXX	Haem domain of CYPXXX
HNQ	2-Hydroxy-1,4-naphthoquinone
HPLC	High-performance liquid chromatography

IPTG	β -d-thiogalactopyranoside
IR	Infrared
k_{cat}	Normalised rate constant
K_{d}	Dissociation constant
K_{H}	Dissociation constant (based on a fit using the Hill function)
KPi	Inorganic phosphate
LB	Luria-Bertani growth medium
m/z	Mass to charge ratio
MALS	Multi-angle light scattering
M_{r}	Molecular mass
MS	Mass spectrometry (or spectrometer)
MSD	Mass spectrometer detector
<i>Mtb</i>	<i>Mycobacterium tuberculosis</i>
MV	Methyl viologen
NADH (NAD ⁺)	Reduced (oxidised) nicotine adenine dinucleotide
NADPH (NADP ⁺)	Reduced (oxidised) nicotine adenine dinucleotide phosphate
NAPQI	<i>N</i> -acetyl- <i>p</i> -benzoquinone imine
NMR	Nuclear magnetic resonance
NPG	<i>N</i> -palmitoyl glycine
NTA	Nitrilotriacetate
OD	Optical density
PCR	Polymerase chain reaction
PDOR	Phthalate dioxygenase reductase
PMS	Phenazine methosulfate
PMSF	Phenylmethylsulfonyl fluoride
PPo	Psi-factor producing oxygenase
P-seph	Phenyl Sepharose
psi(-factor)	Precocious sexual indicator
QELS	Quasi-elastic light scattering
Q-seph	Quaternary amino Sepharose
RCSB PDB	Research collaboratory for structural bioinformatics protein data bank
RDCYPXXX	Reductase domain of CYPXXX
RDX	Royal demolition explosive
R_{h}	Hydrodynamic radius
RING	Really interesting new gene
RNA	Ribonucleic acid

Front Matter

Rz	Reinheit zahl (Reinheit number)
SCE	Standard calomel electrode
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SHE	Standard hydrogen electrode
TB	Terrific broth
TE	Tris-EDTA
UV	Ultraviolet
v/v	Volume to volume
w/v	Weight to volume
YT	Yeast tryptone medium

Abstract

Characterisation of Novel Cytochrome P450 Fusion Systems

Thesis for the degree of Doctor of Philosophy

Jacob W. Robinson

The University of Manchester

September 2010

The biophysical and spectroscopic characterisation of two novel P450 fusion enzymes is reported. The first of these is CYP102A3, which is a fusion of P450 haem and cytochrome P450 reductase (CPR)-like domains and functions as a catalytically self-sufficient fatty acid hydroxylase in its host organism *Bacillus subtilis*. The elucidation of structural aspects of the isolated haem domain of CYP102A3 (HDCYP102A3) is described. This reveals a strong homology between HDCYP102A3 and the haem domain of the related, well studied enzyme CYP102A1 (known as BM3). Examination of the substrate binding and redox properties of HDCYP102A3 reveals variations in substrate selectivity and the influence of substrate binding over the haem-iron redox potential compared to BM3. Of particular note is the apparent cooperative binding profile displayed for some branched chain fatty acid substrates with CYP102A3. The second system characterised is CYP116B1 from *Cupriavidus metallidurans*, a P450 fusion with a reductase domain that resembles phthalate dioxygenase reductase (PDOR). The purification of the intact CYP116B1 enzyme, and also of its isolated haem domain (expressed from the relevant gene section), is optimised and biophysical characterisations are reported. The haem iron redox potential is found to be unusually positive (-85 mV) and the influence of thiocarbamate herbicide substrate binding upon this potential is found to be minimal, unlike the case in CYP102A ϵ with its fatty acid substrates and likely as a consequence of the relatively small degree of shift in haem-iron spin-state towards the high-spin form. From a panel of eight potential substrates for CYP116B1, six were found to stimulate NADPH oxidation, but only two of these were themselves oxidised by the enzyme, with hydroxylated products observable. The genetically dissected reductase domain of CYP116B1 was also expressed and purified, and kinetic studies of the reductase domain revealed a preference for NADPH over NADH coenzyme, and enables comparisons with kinetic features and coenzyme selectivity in other members of the ferredoxin reductase family of enzymes. Collectively, these studies advance our knowledge of the properties of two distinct types of P450-redox partner fusion enzymes, a growing class of enzymes with potential for biotechnological applications.

Declaration

The author hereby declares that the work presented in this thesis has not been submitted in support of an application for any degree, or other qualification, at this or any other university or other institute of learning.

Copyright Statement

- i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University’s policy on presentation of Theses

Acknowledgements

I would like to thank my supervisor Professor Andrew Munro; firstly, for his guidance and encouragement, and secondly, for affording me the opportunity to work in his laboratory with such an excellent group of researchers. I would also like to thank the whole Munro group for their support and acceptance through the course of this project. In particular, I would like to express my gratitude to Drs. Hazel Girvan and Kirsty McLean, whose patient teaching was essential to the success of the majority of laboratory work for this thesis.

Outside of the laboratory, I would like to thank Peter Walch; his friendship, understanding and cuisine, have helped to sustain me through the course of my studies.

Finally, I would like to thank my parents, Lili and Christopher Robinson, for everything they have done for me throughout my life so far. This thesis is dedicated to them for their continued belief in me and their generosity in all things.

Characterisation of Novel Cytochrome P450 Fusion Systems

Chapter 1

Introduction

1. Introduction

1.1. Literature review

1.1.1. What are enzymes?

Enzymes are biological catalysts that allow the reactions required for life to occur on an appropriate timescale under environmental conditions. An extreme example is uroporphyrinogen decarboxylase from the porphyrin biosynthesis pathway which produces prerequisite compounds for essential biological molecules such as haems and chlorophylls. The substrate for this enzyme is predicted to have a half life of 2.3×10^9 years in the absence of enzymatic activity, but can be turned over by the enzyme on the millisecond timescale[1].

Aside from a few examples of enzymes made from RNA or DNA the majority are protein based, being for the most part constructed from a set of 20 naturally occurring amino acid building blocks. The various functional groups present on these amino acids provide both the structural elements of enzyme molecules and in many cases the reactive species required for catalysis. In the latter role amino acids carry reactive groups familiar to synthetic chemistry, chiefly acidic or basic groups and nucleophiles. To expand their chemistry, set some enzymes bind cofactors. These can be anything from single metal ions to large molecules. Cofactors are heavily featured in enzymes carrying out reduction and oxidation (redox) chemistry, where they can act as electron acceptors and donors.

Traditionally, a cofactor that forms a transient interaction with an enzyme is termed a coenzyme while a cofactor that is permanently bound to the enzyme is known as a prosthetic group. Practically, these are somewhat loose terms as some cofactors may have different binding strengths in different enzymes depending on their ligation mode and interactions provided by the protein environment.

1.1.2. Roles of cofactors in enzyme reactions

Some cofactors are effectively carriers for functional groups that are to be transferred to or from the substrate and typically these cofactors contain the functional group in question attached by a bond that can be 'activated' to allow easy transfer. An example

of this type of cofactor is coenzyme A (CoA) which can carry an acyl group. Acetyl CoA is a metabolic intermediate which carries two-carbon units, the product of glycolysis, into the Krebs (tricarboxylic acid) cycle. In the case of CoA the acyl group is attached by a thioester linkage (Figure 1.1) which can be easily broken by reaction with a more powerful nucleophile *e.g.* an OH group[2].

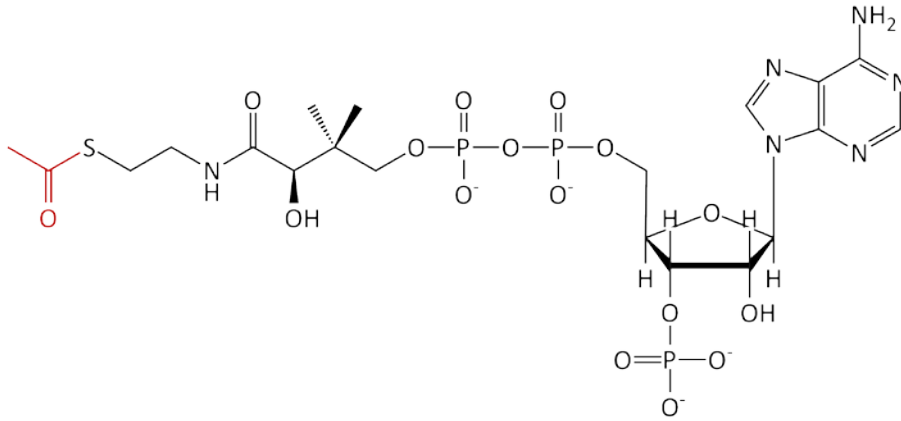


Figure 1.1. Structure of acetyl coenzyme A (CoA). The acetyl group carried via a thioester linkage is shown in red.

Other cofactors, rather than carrying functional groups, are simply involved in the transport of electrons, and these cofactors allow enzymes to carry out redox reactions upon their substrates. Prime examples are the nicotinamide containing cofactors (or coenzymes) nicotinamide adenine dinucleotide (NAD^+) and phosphorylated NAD (NADP^+). Both are produced from nicotinic acid (also known as niacin or vitamin B_3) and are reduced at many stages of metabolism. The reduced forms of these cofactors are represented as NADH and NADPH respectively, which is indicative of the fact that they carry electrons as a hydride (H^-) two electron equivalent. NADH and NADPH are often the source of electrons for redox reactions in enzymes. As each NAD(P)H molecule can only supply two electron equivalents, the coenzyme is turned over during the reaction returning to the oxidised form, usually denoted NAD(P)^+ (Figure 1.2). A fresh NAD(P)H molecule is then required for the next cycle of the reaction.

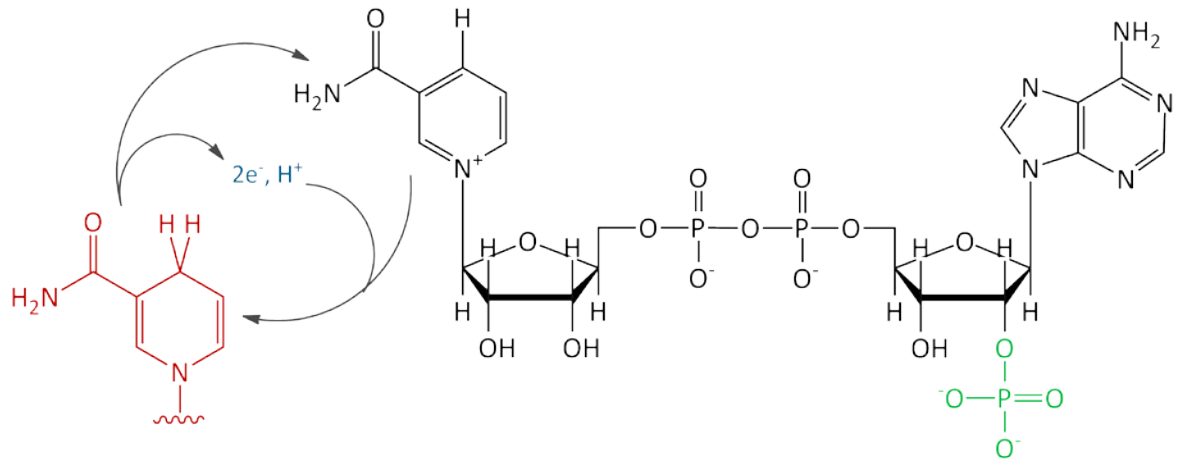


Figure 1.2. Full structure of oxidised nicotinamide adenine dinucleotide phosphate (NADP⁺). The phosphate group (green) on the 2' position of the adenine nucleotide ribose is not present in the related molecule NAD⁺. In red is shown the nicotinamide portion of the molecule in the reduced coenzyme, having picked up a hydride equivalent (2e⁻, H⁺). Note the loss of aromaticity in the nicotinamide ring, which destabilises the reduced form, increasing its reducing power.

Further important types of enzyme cofactor are those based on riboflavin (vitamin B₂), the most important of which are flavin mono-nucleotide (FMN) and flavin adenine dinucleotide (FAD). Both cofactors contain an isoalloxazine ring linked at N10 to the ribosyl chain of a nucleotide moiety. In FMN the nucleotide group consists of the linking ribosyl chain with a terminal phosphate. FAD is identical to FMN with an adenine nucleotide attached by a phosphodiester linkage (Figure 1.3). In these cofactors the isoalloxazine ring is the reaction centre while the nucleotide portion serves as a 'handle' for binding of the cofactor to the protein.

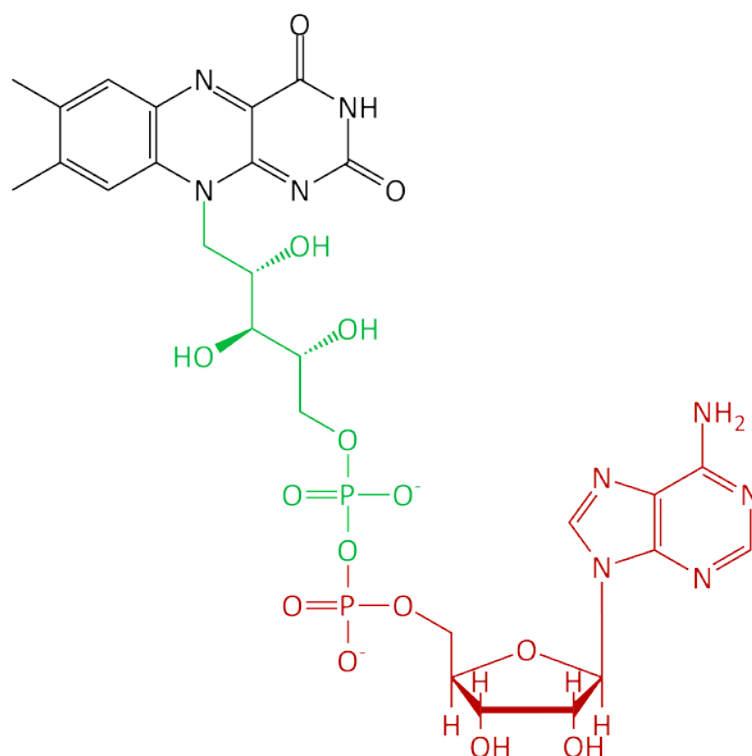


Figure 1.3. Structure of the flavin adenine dinucleotide (FAD) cofactor. The isoalloxazine ring (black) is attached at N10 via a ribosyl chain and a phosphodiester linkage (green) to the adenine nucleotide moiety (red). This group is absent in the related flavin mononucleotide (FMN) cofactor which simply has a single terminal phosphate group after the ribosyl chain.

Flavins are some of the most versatile cofactors, playing diverse roles in many different enzymes[3]. Most commonly, flavins are employed in redox enzymes as electron transfer centres participating in redox reactions either as part of an electron transport chain or as the terminal reducing species. For example, an FMN cofactor is involved in the oxidation of NADH, forming the first step in the electron transport chain of respiratory complex I[4] and flavin as a terminal reductant in flavin monooxygenase enzymes[5]. A flavin cofactor can accept up to two electrons at a time. A one electron reduced flavin is known as the semiquinone form and a two electron reduced flavin is called a hydroquinone. Fully oxidised flavins show absorbances at ~ 360 nm and ~ 450 nm in the UV/visible spectrum and have a yellow colour. The signal at ~ 450 nm is used diagnostically having an absorption coefficient in free flavin of 12.5 (FMN) and 11.3 (FAD) $\text{mM}^{-1} \text{cm}^{-1}$. These extinction coefficients are modified by binding of the flavins to enzymes and also by reduction. The semiquinone can exist as a positive, negative or neutrally charged radical, all of which have a distinctive spectral signal, while the hydroquinone form is spectrally silent across most of the visible spectrum[6].

Alongside flavins, a more recently characterised set of cofactors is the quinone cofactor group. Often derived from the post-translational modification of amino acid residues within the enzyme, these cofactors also participate in redox reactions[7, 8].

Metals, long known to chemists for their catalytic abilities, are the reaction centres in many enzyme cofactors. The most basic cofactor is simply a metal ion ligated by the side chains of amino acid residues. A slightly more complex arrangement is the iron sulfur cluster, of which several arrangements are known. They all consist of a number of iron atoms in complex with a number of inorganic sulfur atoms. A selection of common iron sulfur clusters is shown with appropriate labels showing the nomenclature (*i.e.* xFe-yS where x is the number of iron atoms and y is the number of sulfur atoms) (Figure 1.4).

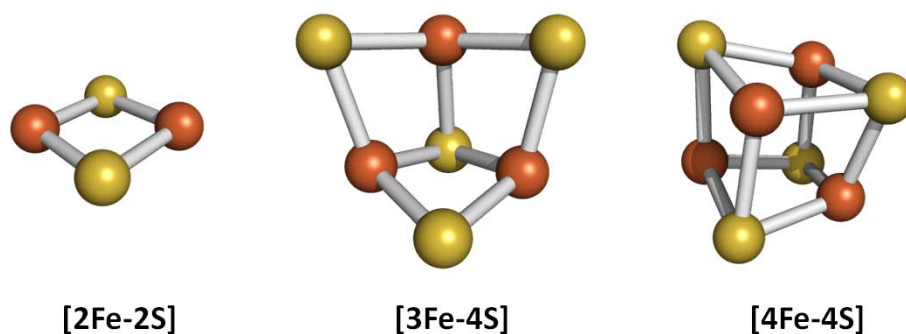


Figure 1.4. A selection of biologically occurring iron-sulfur clusters. Orange spheres represent iron atoms, yellow spheres represent sulfur atoms.

Iron sulfur clusters are commonly bound to enzymes by ligation of the iron atoms by thiolate sidechains of cysteine residues, although alternative ligations do exist. For example, the Rieske-type [2Fe-2S] cluster which is ligated by two cysteine and two histidine residues[9] (Figure 1.5). Inorganic sulfide is also involved in bonds between iron atoms in such clusters.

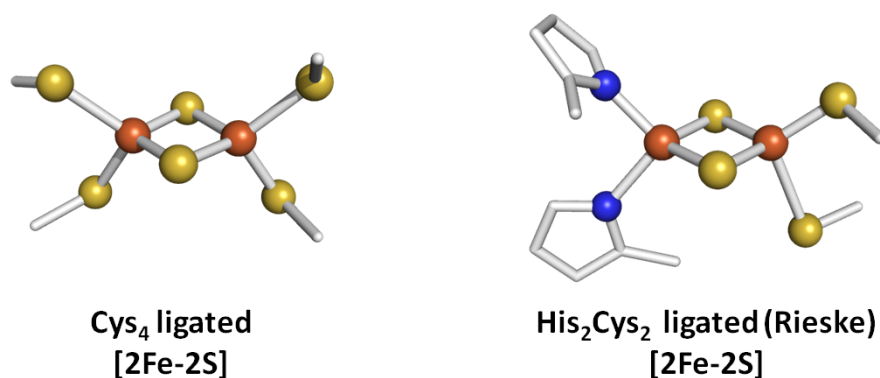


Figure 1.5. Structures of two [2Fe-2S] iron sulfur clusters with the side chains of ligating amino acid residues and inorganic sulfurs between iron atoms. These represent the standard ligation by four cysteine residues (left) and the Rieske-type ligation by two cysteine and two histidine residues (right).

Iron sulfur clusters can participate in redox reactions by acting as electron acceptors and/or donors. The distinct iron atoms within a cluster can exist in different oxidation states from one another at the same time[10]. An example of iron-sulfur proteins is the ferredoxin family of electron transport proteins which shuttle electrons to a variety of other redox enzymes[11, 12].

Probably the best known of all cofactors are haems, which have an iron ion chelated in a porphyrin ring. All haems are based on the same porphine delocalised π -bonding system which is composed of four pyrrole rings linked into a larger aromatic ring (a tetrapyrrole). The orientation of substituent groups on the outside of the porphyrin ring relative to the surrounding residues can finely adjust the electronic properties of the haem[13]. Some common haems are shown in Figure 1.6¹.

¹ By convention haem types are designated by capital letters when considered in isolation and in lower case with italicised letters when bound to a protein (*i.e.* haem A versus haem *a*).

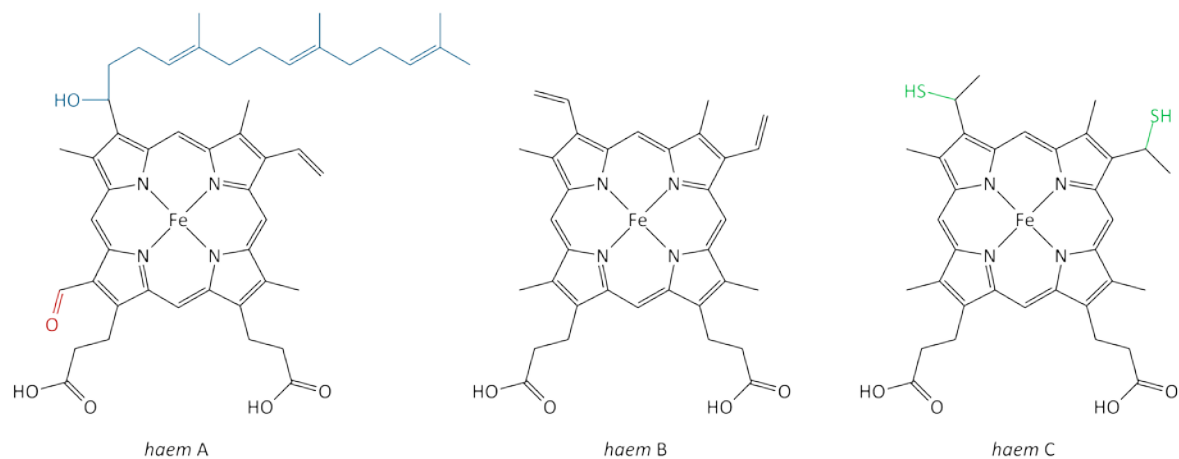


Figure 1.6. Structures of haems A, B and C. Haem B is the most common with methyl groups at position 2,7,12 and 18 and vinyl groups at position 3 and 8. Haem A is a derivative molecule where the methyl group at position 18 is oxidised to a formyl group (red) and the vinyl at position 3 is replaced by a farnesyl group (blue). For haem C, thiol groups from the binding protein are covalently attached across both vinyl groups (green). The latter is an example of a covalently bound haem. Other less common haems do exist with various substituent groups[14].

Haems are best known as oxygen carriers in haemoglobin and myoglobin, the oxygen carrying apparatus in red blood cells and muscle tissue, and the reason for their colouration. The ability to transiently carry oxygen originates in the bonding mode. Transition metal ligands with π -bonding systems (CO, CO₂, O₂, CN, NO *etc.*) can participate in back bonding, when electrons are shared from the metal's d-orbitals into π^* orbitals on the ligand. This occurs in addition to the main σ bond and strengthens the interaction between metal and ligand. Carbon monoxide as a ligand binds iron end on which orients its π^* orbitals in alignment with the iron d-orbitals and enables back bonding to occur. For this reason CO is ligated almost irreversibly to haem (under physiological conditions), which is the origin of the toxicity of CO. In contrast, molecular oxygen (O₂) binds in a bent conformation because the oxygen lone pairs (σ -donors) are at 120° to the oxygen-oxygen bond, which produces poor alignment of the O₂ π^* orbital with the iron d-orbitals. This reduces back bonding and makes O₂ a much more labile ligand than CO.

In addition to simply binding and releasing oxygen as in haemoglobin, a haem may act as a reaction centre for the scission of molecular oxygen, as seen in the oxygenase class of enzymes. In these enzymes either one or both of the oxygen atoms may be incorporated into the substrate, reactions carried out by monooxygenases and dioxygenases respectively.

1.1.3. The Cytochromes P450

The cytochromes P450 (P450s) are a group of enzymes which span the kingdoms of life with examples known in Animals, Plants, Fungi, Bacteria and Archaea[15]. Originally identified in and isolated from mammalian liver microsomes in 1964 by Omura and Sato[16, 17], the following decades of research have revealed a diverse superfamily of enzymes with activity against an equally diverse range of substrates.

All P450s contain a haem *b* cofactor which is the catalytic reaction centre of the enzyme. The haem is coordinated via an axial cysteine thiol/thiolate ligation to the iron and in many cases takes a water molecule as a sixth ligand. P450 enzymes typically have a Soret band absorbance in the visible spectrum at around 420 nm with an absorption coefficient of 90 - 100 mM⁻¹cm⁻¹. When reduced and coordinated to carbon monoxide (CO), at the opposite axial position, in most cases this signal shows a characteristic shift to 450 nm (Figure 1.7). This is the origin of the name P450 which indicates 'Pigment, 450 nm'[16].

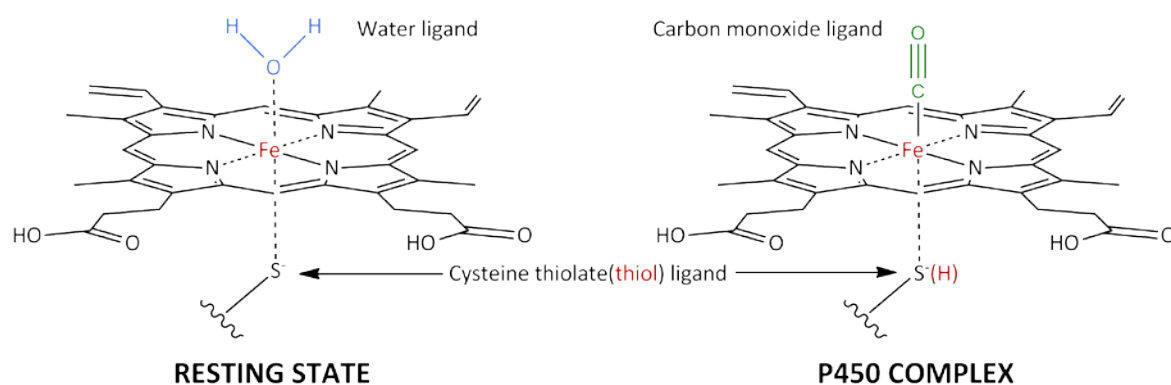


Figure 1.7. Structural representations of the P450 haem environment during its resting state with an axial water ligand (left) and with bound CO (right) Although the carbon monoxide ligand is shown with a formal triple C-O bond, this bond will in fact have a lower order and the C-Fe bond will have a higher order due to back bonding as discussed previously. Under reducing conditions a proportion of the cysteine thiolate ligating the haem may become protonated. This varies from P450 to P450 depending on the effect of the protein environment on the pK_a of the cysteine sidechain.

1.1.4. P450 Nomenclature

As the prevalence of P450 enzymes was revealed, a unified system of nomenclature for P450 enzymes was first introduced in 1987 by Nerbert *et al.* [18] and has been updated and maintained to date by David Nelson (University of Tennessee). This system uses CYP codes to group P450 enzymes related by structure and function *e.g.*

CYP51B1. CYP stands for cytochrome P450 and a number designates the family to which the enzyme belongs. A letter then indicates subfamily and another number labels different enzymes within the subfamily. Initially, P450s were required to share >40 % amino acid sequence identity within a family and >55 % amino acid sequence identity within a subfamily[19]. Later on, the requirement for strict thresholds of sequence identity were relaxed and P450s were grouped by looking at clustering on phylogenetic trees[20]. Ranges were originally assigned for P450s from animals (CYP1-49), lower eukaryotes (CYP51-69), plants (CYP71-99) and bacteria (CYP101 upwards). Due to the number of P450 genes discovered in the years following, further ranges had to be assigned, first in three digits (up to CYP999) and now in four digits (up to CYP9999). To address the rapidly increasing number of P450 families a further level of grouping was introduced. P450 clans group P450 families which cluster on phylogenetic trees and are denominated by the name of the lowest numbered family that belongs to the clan[21]. A maintained list of P450 amino acid sequences with associated CYP codes is found on 'The Cytochrome P450 Homepage' by Dr. David Nelson[22].

1.1.5. Relevance of P450s to Human Health

The human genome is known to contain 57 genes encoding P450 enzymes and many of these play crucial roles in the detoxification and excretion of xenobiotics including drug molecules[23]. The major role of P450 enzymes is in phase I drug metabolism which largely involves increasing the solubility of drug molecules to aid excretion in bile or urine[24]. The most common P450 reaction, hydroxylation, is conducive to this aim as it increases overall polarity of a substrate and therefore its water solubility. In addition, hydroxylation may be an initial step which provides a site for attachment of a further polar group. Though P450s prevent the build up of toxic substances introduced into the body, they can also convert relatively harmless molecules into toxins. An example is CYP2E1 which carries out the first step in the detoxification of many xenobiotic molecules. One such molecule is the common analgesic drug paracetamol (also known as acetaminophen) for which CYP2E1 is the first step in a minor detoxification pathway. CYP2E1 hydroxylates the amide nitrogen of the paracetamol molecule allowing a rearrangement to form a reactive quinone intermediate; *N*-acetyl-*p*-benzoquinone imine (NAPQI). NAPQI is then conjugated to

the polar molecule glutathione (GSH) by the enzyme glutathione S-transferase (GST). In cases of paracetamol overdose or glutathione deficiency, excess NAPQI can accumulate and react with cellular proteins. This is thought to be the origin of liver toxicity associated with paracetamol overdose and cases of paracetamol use in individuals with depleted GSH and/or increased CYP2E1 expression[25, 26].

As well as degradation pathways, P450s are also involved in biosynthetic processes including the manufacture of steroid hormones[27]. For example, CYP19A1 is an aromatase enzyme, found in human oestrogenic cells, that performs a key reaction in the biosynthesis of oestrogen[28]. CYP51 and related families are 14 α -sterol demethylases found in organisms from all kingdoms of life including humans. The reaction they catalyse forms part of the eukaryotic sterol biosynthesis pathway[29]. Interruption of such essential metabolic routes, either by the presence of mutant enzymes or inhibition of functioning enzymes, can cause an array of disorders. For example, the presence of mutant forms of CYP21A2, a P450 involved in human adrenal hormone biosynthesis, accounts for more than 90 % of all cases of congenital adrenal hyperplasia (CAH)[27]. This group of conditions leads to a deficiency of the hormone cortisol which can cause problems during foetal development[30].

Both their catabolic and anabolic roles in humans make the study of P450s an important area for our understanding of the fate of drugs once they are introduced into the body and how to mitigate any potentially unwanted effects arising from these interventions[31, 32].

P450s can also present promising drug targets in human pathogens such as *Mycobacterium tuberculosis* (*Mtb*). The *Mtb* genome project revealed that the organism has 20 genes coding for P450 enzymes, several of which have been shown to be essential for either viability or pathogenicity of the organism[33-35]. For example, *CYP51B1*, a gene encoding a sterol demethylase enzyme of the CYP51 family, is potentially important to the organism for metabolising host sterols during infection and therefore also as a drug target[36, 37]. *CYP121A1*, a gene found to be essential for viability of *Mtb*, encodes the CYP121A1 enzyme which has a demonstrable activity against a cyclodipeptide substrate, a possible step in synthesis of a more complex toxin or effector of the host metabolism[38, 39].

1.1.6. The Reactions Catalysed by P450s

The most commonly observed P450 reaction is monooxygenation, either hydroxylation or epoxidation, at a carbon centre (Figure 1.8). However, many other, less common reactions have been identified which range from oxidations and dealkylations at a heteroatom (S, N, P *etc.*)[40], to a range of complex reactions. These include single electron reductions, desaturations and ring modifications, all of which are often observed as a minor reaction alongside the traditional oxidation (reviews [41, 42]). Many of these are reactions that can be difficult to achieve by synthetic chemistry and, if feasible, certainly occur without the regio- and stereo-selectivity of which the enzymes are capable.

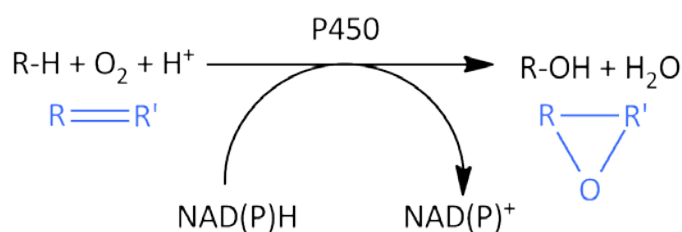


Figure 1.8. The classical P450 reaction: oxidation either as hydroxylation of a carbon centre or epoxidation across an unsaturated carbon-carbon bond (in blue).

In the resting state a water ligand is coordinated to the distal axial position of the P450 haem iron and this usually favours a low-spin configuration of the haem iron d-electrons. When the substrate binds in the active site it displaces the water leaving a pentacoordinate haem iron which changes the favoured electron configuration to high-spin (Figure 1.9). This change in spin state also causes a shift in the Soret band associated with the haem iron towards a shorter wavelength, known as a type I shift. Equally, when strong field ligands are bound to the haem iron, either directly or through a water network, a low-spin complex is favoured and this causes the Soret band to shift to a longer wavelength, known as a type II shift. This type of ligand, if it is bound tightly, can act as an inhibitor of the enzyme. These spectral signals were identified in liver microsome P450s treated with substrates and haem ligands[43].

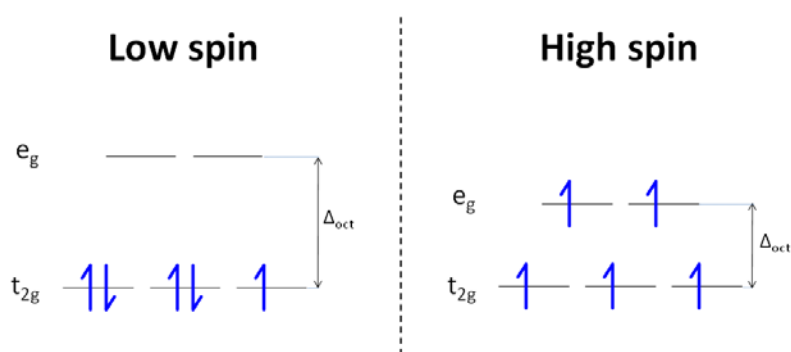


Figure 1.9. Diagram representing the d-electron configuration of ferric iron (Fe³⁺), showing the distribution of electrons among the two groups of d orbitals (denoted e_g and t_{2g} from the space groups they occupy) in high spin and low spin states. Electronegativity of ligands to the iron can raise or lower the octahedral splitting energy (Δ_{oct}) between the orbital groups. When this is less than the electron pairing energy a high-spin state is favoured; when it is greater a low-spin state is favoured.

This high-spin change causes a positive shift in the redox potential of the haem, facilitating a one electron reduction of the iron to the ferrous (Fe²⁺) oxidation state[44]. This is a form of substrate regulation of the enzyme which prevents futile cycling of electrons through the system (and consumption of NAD(P)H) in the absence of substrate[45]. Once the haem iron is reduced, O₂ can become coordinated at the axial position and reduced to a superoxide oxidation state (O₂⁻) with the iron returning to a ferric (Fe³⁺) state. A further one electron reduction then produces a haem peroxide species which can readily capture one proton to form a hydroperoxide, this is known as Compound 0. The hydroperoxide can then attract a second proton allowing the first oxygen atom to leave as a water molecule and yield a

ferryl-oxo ($\text{Fe}^{4+}=\text{O}$) intermediate known as Compound I. The designations compound 0 and compound I are based on equivalent species identified in peroxidases[46]. Both are highly reactive and transient radical species and both have been suggested as the terminal oxidant that attacks the substrate in P450 enzymes[47]. Current evidence points to Compound I as the most likely candidate for the oxidising species in P450 hydroxylation reactions [48] and a mechanism known as radical rebound has been proposed[49]. This mechanism has been demonstrated in model systems[50, 51] and involves the abstraction of a proton from the substrate by compound I, leading to the formation of a ferryl-hydroxy ($\text{Fe}^{4+}\text{-OH}$) species (Compound II) and a substrate radical. The substrate radical then attacks the $\text{Fe}^{4+}\text{-OH}$ haem forming a new bond to the OH group and restoring the haem iron to the Fe^{3+} state[52] (Figure 1.11). The oxidised product diffuses out of the active site and is replaced with a water molecule to return the enzyme to its resting state. This constitutes the generally accepted model of the P450 reaction cycle[44, 53] (Figure 1.10).

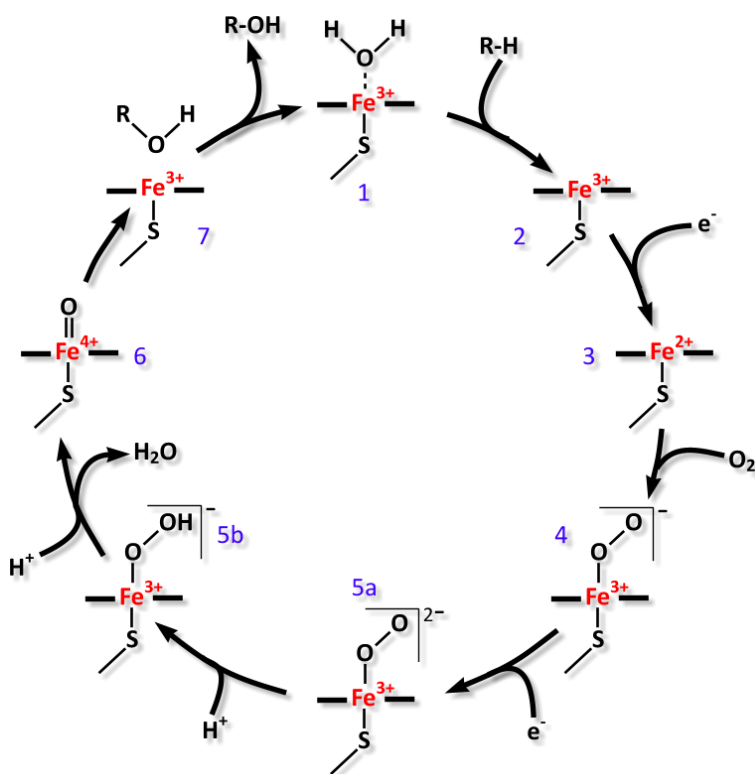


Figure 1.10. Schematic diagram of the catalytic cycle of a cytochrome P450. 1-2, Substrate diffuses into the active site, near to the haem, displacing the water ligand and promoting a high-spin state. 2-3, haem iron undergoes a one electron reduction to Fe^{2+} . 3-4, Molecular oxygen (O_2) is bound to haem iron and reduced to produce Fe^{3+} -superoxide. 4-5a, Superoxy-haem undergoes a further one electron reduction to form an Fe^{3+} -peroxy species. 5a-5b, Protonation to hydroperoxide (Compound 0) occurs. 5b-6, Scission of the O-O bond to produce a water molecule and an ferryl-oxo (Fe^{4+}) species (Compound 1). 6-7, Ferryl-oxo species attacks the substrate via the proposed radical rebound mechanism (Figure 1.11) to leave hydroxylated substrate (R-OH) and pentacoordinate Fe^{3+} haem. 7-1, Product diffuses out of the active site and is replaced by a water ligand to restore the enzyme to its resting state. Based upon a figure from Denisov et al. 2005[44]

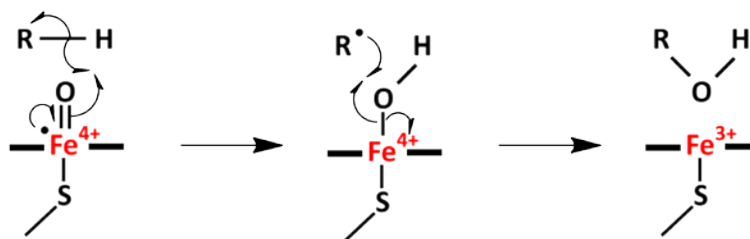


Figure 1.11. Proposed radical rebound mechanism describing the terminal oxidation event in the P450 catalytic cycle. Compound I (left) extracts a proton from the substrate (R-H) by heterolytic cleavage of the R-H bond to produce a substrate radical (centre). The substrate radical attacks (rebounds on to) the resultant ferryl-OH (Compound II) forming a bond to the OH group and returning the haem iron to the Fe^{3+} resting state (right). Based upon an illustration by Groves and Subramanian[51]

Although hydroxylation is the most commonly observed P450 reaction, many other less common reactions have been identified[41]. These include the *N*-dealkylation of amines which has been shown to proceed by a single electron oxidation at the N atom followed by proton abstraction and hydroxylation at an alkyl carbon. This hydroxylated product would decompose to leave an *N*-dealkylated substrate and an aldehyde.

1.1.7. Structure of P450s

While P450s in the same family may show a high degree of amino acid sequence identity (>40%), the larger superfamily shows a wide variation aside from some highly conserved motifs such as the haem binding sequence (see Figure 1.12).

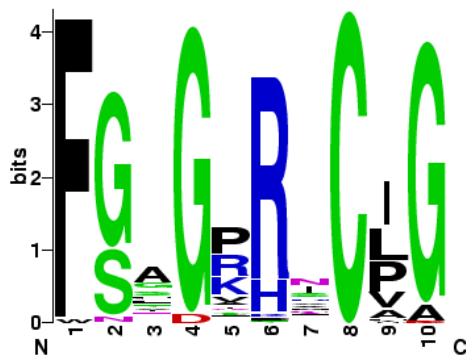


Figure 1.12. Logo showing the consensus amino acid sequence of the P450 haem binding motif derived from alignments of 868 sequences containing the motif. The relative height of the amino acid letter codes indicates their frequency of occurrence in this motif. The absolutely conserved cysteine residue (8) provides the thiol/thiolate ligand to the haem. The logo was created using WebLogo through the ExPASy Prosite tool.

Despite this variation, all P450s share a common fold with shared secondary and tertiary structural features visible in their structures.

The first P450 enzyme to be structurally characterised was CYP101A1 (P450cam), a camphor hydroxylase from *Pseudomonas putida*. This was the first bacterial P450 to be discovered[54, 55] and, as a soluble protein, presented a major advantage for characterisation over exclusively membrane-associated, mammalian P450s. The first crystal structure of P450cam was published in 1985[56] (Figure 1.13a).

Since this many other P450 enzymes have been structurally characterised and the common structural features of P450 enzymes have been revealed (Figure 1.13).

The P450 fold comprises two subdomains, one composed mainly of α -helix secondary structure and another largely composed of β -sheet. In the α subdomain the number and orientation of helices varies, but an invariant feature is the I-helix, a long helix spanning the subdomain and passing close to the haem group.

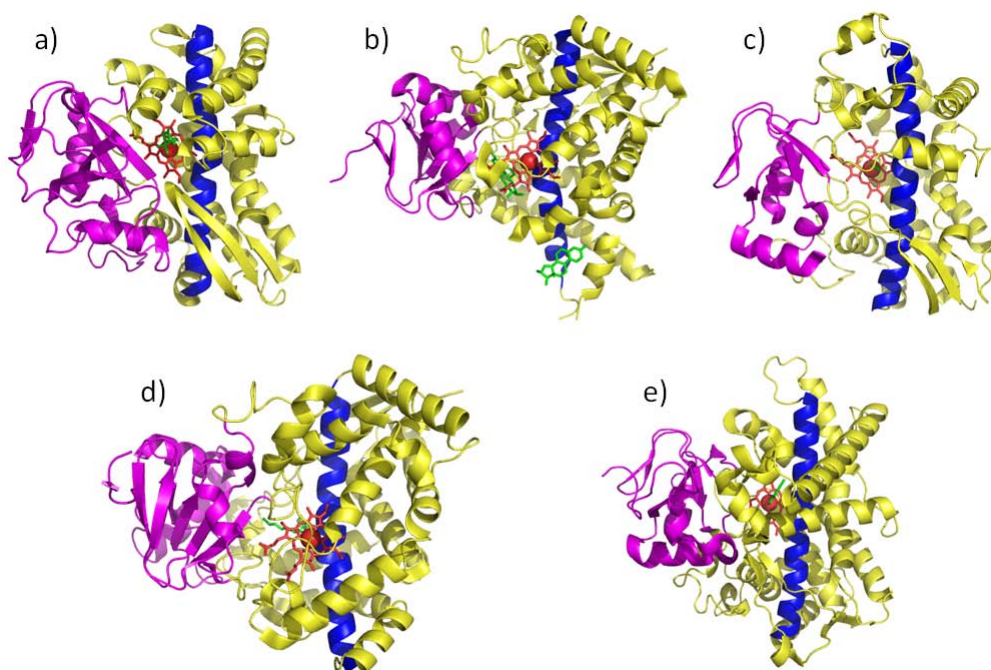


Figure 1.13 Structures of cytochrome P450 enzymes from various organisms showing structural similarity. Each enzyme comprises a minor subdomain with extensive β -sheet structure (magenta) and a major subdomain of mostly α -helix structure (yellow). The latter contains the haem *b* cofactor (red) and the I-helix (blue) a long α -helix spanning the enzyme and running adjacent to the haem. The enzymes are as follows a) CYP101A1 camphor hydroxylase from *Pseudomonas putida*[57]. b) CYP51 sterol demethylase from *Mycobacterium tuberculosis*[58]. c) CYP119A1 enzyme with fatty acid hydroxylase activity from *Sulfolobus solfataricus*[59]. d) CYP74A1 allene oxide synthase from *Parthenum argentatum*[60]. e) CYP2A6 drug metabolising enzyme from *Homo sapiens*[61].

1.1.8. Redox Partners

Electrons needed to reduce the haem reaction centre in P450s are generally provided by redox partners, other enzymes which accept electrons from the biological electron carrier reduced nicotinamide adenine dinucleotide (NADH) and/or its phosphorylated form (NADPH)[62]. These were originally thought to be divided between two classes. The class I reductases are homologous to ferredoxin (Fd) and ferredoxin reductases (FdR) with iron sulphur cluster (Fd) and FAD (FdR) cofactors, and a NAD(P)H binding site in the latter[53]. The class II reductases consist of a Cytochrome P450 reductase (CPR) enzyme containing flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) cofactors, and again a NAD(P)H binding site[53]. Analysis of amino acid sequences has predicted that CPR genes are the result

of a fusion event between genes coding for flavodoxin and flavodoxin (or ferredoxin) reductase-like proteins at some point during the evolution of the enzyme[63].

It was thought, based on the evidence, that these two classes represented the extent of P450 redox partner systems; the class I system representing a prokaryotic model and the class II occurring exclusively in eukaryotic species. It was shown that mammalian mitochondrial P450s employ a class I-like redox partner system consisting of a [2Fe-2S]-containing ferredoxin (adrenodoxin) and FAD-binding adrenodoxin reductase[64]. This discovery lent support to the hypothesis of class I/class II divide, based on the evolutionary origin of the mitochondrion from *prokaryota*. The discovery of CYP102A1, known as P450-BM3 (BM3), in the genome of bacterium *Bacillus megaterium* went against the pattern, as it was found to employ a CPR-like diflavin reductase[65]. BM3 was also the first example of a P450-redox partner fusion enzyme, that is to say, an enzyme where the P450 reaction unit and its redox partner(s) are expressed as a single protein chain[40] (discussed in more detail in the next section).

There is also an example of a P450 enzyme that does not require a redox partner. CYP55A1, known as P450_{nor}, from *Fusarium oxysporum*, carries out the reduction of nitrite through nitric oxide to N₂O. This P450 enzyme is able to abstract electron equivalents directly from NAD(P)H to directly reduce its haem centre[66].

1.1.9. The History of Fusion P450s

As mentioned above, the first P450 fusion enzyme to be characterised was CYP102A1, a fatty acid hydroxylase from *Bacillus megaterium* (also known as BM3) discovered in 1986 by Narhi and Fulco[65]. It consists of a 54 kDa P450 domain covalently linked to a 64 kDa (Class II) CPR-like diflavin reductase domain to form a 118 kDa enzyme with both P450 and CPR activities[67]. The enzyme was described, in this initial work, as a “Catalytically self-sufficient...Cytochrome P450 monooxygenase” indicating that the whole redox apparatus from NAD(P)H oxidation to oxygen scission and substrate reaction were contained in one protein chain[68]. At the point of its discovery, this enzyme was also unusual in that the fused redox partner for this bacterial P450 was homologous to CPR, the Class II redox partner, at the time thought to be exclusive to mammalian type P450 enzymes[65].

BM3 has the fastest recorded rate of any P450 at 17 000 min⁻¹ with the substrate arachidonic acid[69]. This is thought to be due to the fused nature of BM3, the covalent attachment of its redox partner producing a dramatic increase in the rate of electron transfer through the system during turnover compared to P450s with separate redox partners[70]. The fast reaction rate combined with the ease of expression and purification of the soluble bacterial protein has made BM3 a useful and enduring model of human P450s[71].

Sedimentation experiments had suggested that BM3 existed in solution as a mixture of monomer, dimer, trimer and some higher order species[72]. Mutagenesis was used to produce two inactive enzymes, one with a defective haem domain and one with a defective reductase domain, both mutants being unable to catalyse substrate hydroxylation individually. When combined in solution fatty acid hydroxylase activity was reconstituted[73]. These experiments confirmed what had been suggested by dilution studies, that BM3 is only fully catalytically active as a dimer, with electron transfer occurring between the FMN of one monomer and the haem of the other[74].

Around 1996 details of a fusion P450 with the same domain structure (P450-CPR fusion) and catalytic function as BM3, from a fungal organism, were published. CYP505A1, from *Fusarium oxysporum*, (known as P450foxy) shares 36 % sequence identity with BM3 and is a fatty acid hydroxylase enzyme[75]. Evidence shows that, in *F. oxysporum*, CYP505A1 is expressed as a soluble unit equivalent to BM3 and a post-translational modification causes it to become membrane bound[76]. Turnover studies showed that CYP505A1 catalyses its reaction at a comparable rate to BM3 with which it also shares the preference for hydroxylation at the ω -2 position[77].

In 2001 two additional members of the CYP102 family, CYP102A2 and CYP102A3, were identified from the genome of *Bacillus subtilis*[78]. Both of these enzymes share >50% amino acid sequence identity with P450 BM3 and are also fatty acid hydroxylase enzymes, although with altered substrate preferences and regioselectivities[79, 80]. Since the discovery of these two enzymes several other members of the CYP102 family have been discovered, the majority of which are in the CYP102A subfamily[22]. All of the CYP102A enzymes appear, from amino acid sequence alignments, to be fusion systems of P450 and CPR-like reductase domains with varying levels of sequence identity with BM3 (Table 1.1). A phylogenetic tree of

the sequences of enzymes of the CYP102A subfamily seems to show the sequences clustering into three subgroups (Figure 1.14).

CYP102 Enzyme	Organism	% Sequence Identity with A1
A1 (BM3)	<i>B. megaterium</i>	100
A2	<i>B. subtilis</i>	59
A3	<i>B. subtilis</i>	58
A4	<i>B. anthracis</i>	60
A5	<i>B. cereus</i>	61
A6	<i>Bradyrhizobium japonicum</i>	47
A7	<i>B. licheniformis</i>	59
A8	<i>B. thuringiensis serovar konkukian</i>	60
A9	<i>B. weihenstephanesis</i>	61
A10	<i>Erythrobacter litroalis</i>	46
A11	<i>Erythrobacter sp.</i>	47
A12	<i>Rhodopseudomonas palustris</i>	47
A13	(uncultured soil bacterium)	46

Table 1.1. List of enzymes in the CYP102A subfamily with host organism and percentage amino acid sequence identity to CYP102A1 (BM3). The enzymes are those with sequences listed in the comprehensive list at David Nelson's P450 homepage[81]. Sequence identity scores are from an alignment carried out using ClustalW2[82].

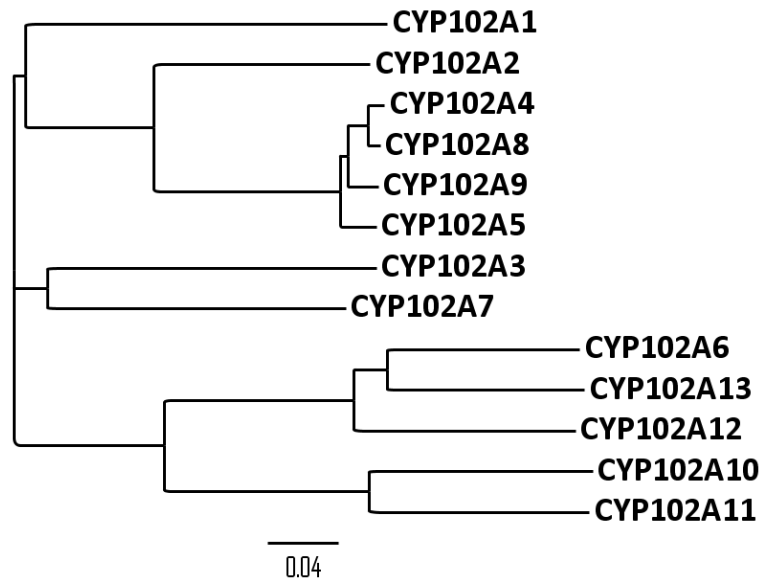


Figure 1.14. Phylogenetic tree showing all presently sequenced members of the CYP102A subfamily. Figure produced using FigTree Tree figure drawing tool by Andrew Rambaut.

In addition to the CYP102A type, several other P450-redox partner fusions have been identified and all are shown schematically in Figure 1.15. The first of these fusions is a P450 haem domain fused to a domain resembling phthalate dioxygenase reductase (PDOR). At least three P450s with this fusion arrangement are known, all of which are members of the CYP116B subfamily (more in Chapters 4 & 5). The next fusion is an N-terminal flavodoxin fused to a C-terminal P450, the reverse arrangement to the other fusions. This type of fusion is found in CYP177A1 (XplA) (more about which, in section 4.1). The last configuration is a P450 covalently linked to a ferredoxin-like domain containing a [3Fe-4S] cluster. This fusion is found in the CYP51B1 enzyme from *Methylococcus capsulatis* (known as McCYP51FX)[83].

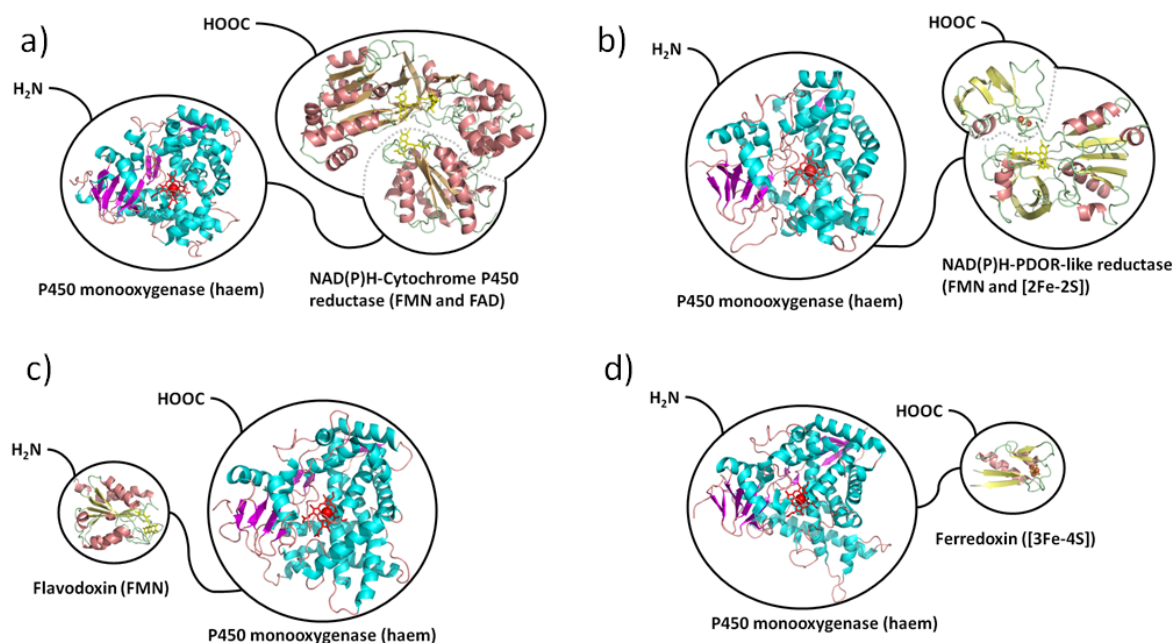


Figure 1.15. Schematic representations of currently confirmed P450 fusions constructed from structures in the RCSB protein data bank (accession numbers in brackets). a) P450-CPR fusion exemplified by BM3 (CYP102A1) represented by the BM3 haem domain (1FAG) and CPR from *Rattus norvegicus* (1AMO). b) P450 fusion to a phthalate dioxygenase reductase (PDOR)-like domain represented by haem domain of P450EryF (CYP107A1) and PDOR from *Pseudomonas cepacia* (2PIA). c) N-terminal flavodoxin-P450 fusion such as XplA (CYP177A1) represented by human CYP2D6 (2F9Q) and flavodoxin from *Escherichia coli* (1AG9) d) P450 fusion with a [3Fe-4S] ferredoxin, represented by CYP51A1 from *Mycobacterium tuberculosis* and [3Fe-4S] ferredoxin from *Pyrococcus furiosus* (1SJ1). Adapted from a figure by Munro *et al.*[84].

1.1.10. Engineering BM3

Since work began on BM3, many insights into the structure-mechanism relationship have been discovered. Trp-96 is a residue conserved throughout the CYP102 family that has been shown, structurally and by mutagenesis in BM3, to be involved in hydrogen bonding to haem propionate groups thus playing an important role in haem binding[85, 86]. Phe-87 has a role in controlling the regioselectivity of substrate hydroxylation by the enzyme. When substrate is bound it undergoes a change in orientation to shield the ω carbon of the substrate from the haem, preventing it from being hydroxylated[87]. Arg-47 and Tyr-51 are a pair of residues shown to interact with the carboxyl terminal of fatty acid substrates, the former of which contributes the majority of the interaction. Mutation of these residues leads to a drop in catalytic rate against substrate arachidonic acid[69]. Phe-42 is thought to form a hydrophobic cap to the active site cavity. Mutation of this residue decreased the observed binding affinity of substrate which is postulated to be due to the presence of water in the

active site increasing polarity and weakening Arg-47 guanidinium to fatty acid carboxylate interactions[69].

These mechanistic discoveries fuelled efforts to alter the reactivity and substrate specificity of the enzyme. The methods used comprise rational and random mutagenesis with selection for a desired activity (directed evolution) in the amino acid sequences of the enzymes. A combination of the two approaches has also been used with some success. The earliest work was on Phe-87, the mutation of which was shown to have an effect upon the regioselectivity of the BM3 reaction. In the first studies it was concluded that the F87A mutant strongly favours hydroxylation at the ω terminal for lauric and myristic acids where in the wild type reaction no hydroxylation occurs[88]. However, this work has been called into question, and it appears to be the case that F87 mutation actually moves hydroxylation further away from the ω -terminus[89].

Further work was aimed at altering substrate specificity and produced mutants of BM3 that show increased binding and catalytic activity against fatty acid substrates with shorter chain lengths (C_4 - C_{10}) compared to the wild type which favours longer chain lengths ($>C_{12}$). This was accomplished with a rational approach by adding residues able to form hydrogen bonds with the carboxylate group of the substrate further into the substrate binding channel. Two mutants of BM3 (L181K and the double mutant L181K/L75T) were produced that had improved turnover against butyric acid (C_4) and hexanoic acid (C_6) respectively. Additionally, if the existing residues thought to fulfil this role (R47 and Y51) were also removed, this had a deleterious effect on substrate binding and catalytic rate regardless of substrate chain length[90].

Much work has been published on the discovery and development of BM3 activities towards various organic compounds from simple alkanes to complex polycyclic aromatic compounds[91-94]. Another avenue in the engineering of BM3 is adaptation of the enzyme to produce drug metabolites. These molecules can be expensive to produce by synthetic chemistry and can be revealed to be effective drugs in themselves[95]. Two studies have produced chimeric enzymes in an attempt to create new activities and improve stability of BM3 (*e.g.* to temperature), in order to allow it to be more easily incorporated into industrial processes such as synthesis of fine chemicals. These chimeras were created by 'mixing and matching' strategically

chosen parts of the CYP102A1 gene with parts of the CYP102A2 and A3 genes. Some of the chimeric enzymes created were found to retain activities comparable to the parent enzymes while having increased thermostability[96, 97].

1.2. Discussion of theory behind the techniques used

The work in this thesis relies on many techniques to study P450 fusion enzymes. The theory underpinning a number of key techniques employed, is described in this section. The practical aspects of these techniques (*i.e.* the actual protocols followed), are described in chapter 2.

1.2.1. Enzyme Kinetic Studies

Assays of enzyme kinetics require that, during turnover, a change occurs, the rate of which can be monitored. Commonly this could be a change in absorbance where either a substrate or product includes a chromophore, with a known molar absorbance coefficient, which is destroyed or created respectively in the enzyme reaction. An example would be ferrochelatase enzymes which incorporate ferrous iron into porphyrins to produce haems which have a strong Soret absorbance around 400-450 nm[98]. In this enzyme reaction, either the consumption of porphyrin or the production of haem could be measured. In other cases, where neither substrates nor products contain a chromophore, changes in absorbance due to a coenzyme can be monitored. This is true in the case of enzymes that use the biological reducing agents NADH or NADPH which, in their oxidised forms, both absorb in the visible/near UV region at 340 nm. When these coenzymes are oxidised they no longer absorb and therefore the rate at which they are consumed in the reaction can be monitored by UV/Visible spectroscopy. From measurements such as these, a value of the observed rate or velocity (v) can be obtained and manipulated to give a value in terms of concentration change over time. When corrected for the enzyme concentration and if the stoichiometry of the reaction is known, this gives a value of the rate of turnover of the enzyme. Since in many enzyme reactions the rate is dependent on the concentration of substrate present, assays are often repeated at a range of substrate concentrations to characterise this dependency and determine (or project) the rate at saturation. This is achieved by fitting the data acquired to one of the many steady-state approximation models available. For a simple first order dependency a

hyperbolic relationship between rate and substrate concentration is expected. This situation can be described by the Michaelis-Menten model which is a steady-state approximation obeyed by many enzyme reactions[99].

$$v = \frac{V_{max}[S]}{K_M + [S]}$$

Equation 1.1 The equation describing the Michaelis-Menten model of enzyme kinetics. Where V_{max} is the maximal rate achieved at saturation with substrate, $[S]$ is the substrate concentration used to generate a rate v , and K_m is the apparent substrate concentration where $v = 50\%$ of V_{max} .

In some cases a biphasic or sigmoidal dependency is observed and this can arise from cooperativity of substrate binding. In these cases data can often be approximated to the Hill model.

$$v = \frac{V_{max}[S]^n}{K_H^n + [S]^n}$$

Equation 1.2 The equation describing the Hill model which can be applied to enzyme reactions exhibiting cooperativity. Here (for enzyme reactions), the terminologies are similar to those for the Michaelis-Menten equation (Equation 1.1), with K_H replacing K_M as substrate concentration at 50 % V_{max} and “n” as a measure of cooperativity, that can reflect the number of substrate molecules involved in the reaction.

In cases where the substrate is bound to the enzyme very tightly (*i.e.* with a dissociation constant close to or lower than the enzyme concentration in solution) the data may be better fitted to a quadratic function also known as a Morrison model. This model takes into account the enzyme concentration present in the assay whereas the other models described above assume enzyme concentration to be negligible compared to that of the substrate.

$$v = \left(\frac{V_{max}}{2[E]} [E] + [S] + K \right) - \sqrt{([S] + [E] + K)^2 - 4[S][E]}$$

Equation 1.3 The equation describing the Morrison model of enzyme kinetics. In this equation, $[E]$ describes the enzyme concentration used, and K is the apparent substrate concentration at 50% V_{max} .

Each of these models provides a theoretical maximal rate (at substrate saturation) and a constant K which is a measure of substrate affinity in the enzyme active site.

Additionally, the Hill model provides a value (n) known as the Hill number which, in the case of cooperative substrate binding, shows the degree of cooperativity involved.

The kinetics of P450 reactions can be monitored easily as in most cases NAD(P)H coenzyme is consumed in their course. As long as NAD(P)H oxidation (*i.e.* hydride transfer to the reductase partner) is not the rate limiting step in the overall reaction, this allows for easy monitoring of the P450 rate, as the oxidation of NAD(P)H leads to a decrease in the signal at 340 nm in the UV/visible absorbance spectrum. This allows rates of reaction to be derived and compared for different P450 reactions.

1.2.2. Ligand Binding

When the binding of a ligand, such as a substrate or inhibitor, to an enzyme can be monitored, (*e.g.* spectrophotometrically) this can be used to quantify the affinity of the enzyme for that ligand. In the case of P450 enzymes, as detailed previously (see section 1.1.6), the binding of substrate and inhibitor ligands cause discernable spectral shifts. These shifts are observed in the Soret band arising from absorbance, in the visible region, by the enzymes' bound haem group.

A binding titration is an experiment to investigate binding of a particular ligand to an enzyme such as a P450. The UV/visible absorbance spectrum of a sample of the enzyme is recorded and then, to the same sample, small-volume additions of the ligand in question are made from a concentrated stock solution. After each addition of ligand the spectrum is recorded again to monitor changes in absorbance associated with ligand binding. In this way the concentration of ligand is increased in small steps and the associated changes in the absorbance spectrum are recorded. To highlight these changes, especially when shifts in absorbance wavelength occur, the spectrum recorded before the addition of ligand is subtracted from each subsequent spectrum to produce difference spectra. In the difference spectra, increases or decreases in absorbance, relative to the ligand-free sample, are obvious. In the case of a shift from one wavelength to another, as in type I and II Soret shifts for P450 enzymes, the differences in absorbance due to the two species are subtracted to produce $\Delta\Delta$ Absorbance values (the difference in the change in absorbance at two wavelengths). These values can then be plotted against the concentration of the ligand to produce a binding curve showing the dependence of binding upon ligand

concentration. Binding curves can be fitted using similar equations as for enzyme kinetic curves (assuming reversible binding, see section 1.2.1) with maximal absorbance (A_{\max}) substituted for V_{\max} . The equivalent of the Michaelis-Menten equation is a hyperbolic one-site binding equation (Equation 1.4) and the Hill (Equation 1.5) and Morrison (Equation 1.6) equations can also be adapted for cases of apparent cooperativity and tight binding respectively.

$$A = \frac{A_{\max}[L]}{K_d + [L]}$$

Equation 1.4. Hyperbolic, one-site binding equation. Where A is absorbance, A_{\max} is maximal absorbance, $[L]$ is the ligand concentration and K_d is the dissociation constant.

$$A = \frac{A_{\max}[L]^n}{K_H^n + [L]^n}$$

Equation 1.5. Hill equation for ligand binding. Where A is absorbance, A_{\max} is maximal absorbance, $[L]$ is the ligand concentration, K_H is the apparent dissociation constant and n is the Hill number. This equation is used in cases where there is apparent cooperativity in the binding of a ligand.

$$A = \left(\frac{A_{\max}}{2[E]} [E] + [L] + K_d \right) - \sqrt{([L] + [E] + K_d)^2 - 4[L][E]}$$

Equation 1.6. Quadratic (Morrison) equation for ligand binding. Where A is absorbance, A_{\max} is maximal absorbance, $[E]$ is the enzyme concentration, $[L]$ is the ligand concentration and K_d is the dissociation constant. This equation is used in cases where binding is very tight ($K_d \approx [E]$).

Each of these equations provides a value of K_d , the dissociation constant (or apparent dissociation constant in the case of the Hill model), which is the ligand concentration at which A is half A_{\max} *i.e.* 50 % of binding sites are saturated with ligand. K_d can be viewed as a measure of the affinity of the enzyme for the ligand in question, the lower the value of K_d , the higher the binding affinity.

For P450 enzymes, K_d values for different ligands can be compared for the same enzyme to determine which substrates or inhibitors bind most efficiently. Equally, K_d values for the same ligand can be compared for different P450 enzymes to show differences in substrate selectivity or inhibitor binding between evolutionarily related enzymes or different mutant forms of the same enzyme.

1.2.3. Resonance Raman Spectroscopy

Raman spectroscopy is a technique for probing intramolecular motions, chiefly bond vibrations and rotations. It is a derivative of infra-red (IR) spectroscopy which measures the same phenomena by observing the absorbance of light in the infra-red region of the spectrum[100]. While IR spectroscopy involves collecting light scattered elastically by a molecule (Rayleigh scattering), Raman spectroscopy uses light that is scattered inelastically (Raman scattering) (Figure 1.16). IR and Raman spectroscopy can show different vibrational modes of a molecule of interest because in some cases modes that are IR silent (*i.e.* cause no elastic scattering) are detectable by Raman and *vice versa*. Because the majority of incident light is scattered elastically, Raman is a somewhat insensitive technique. This insensitivity can make Raman spectroscopy impractical for studying biological systems, with very high sample concentrations required to produce a useful signal.

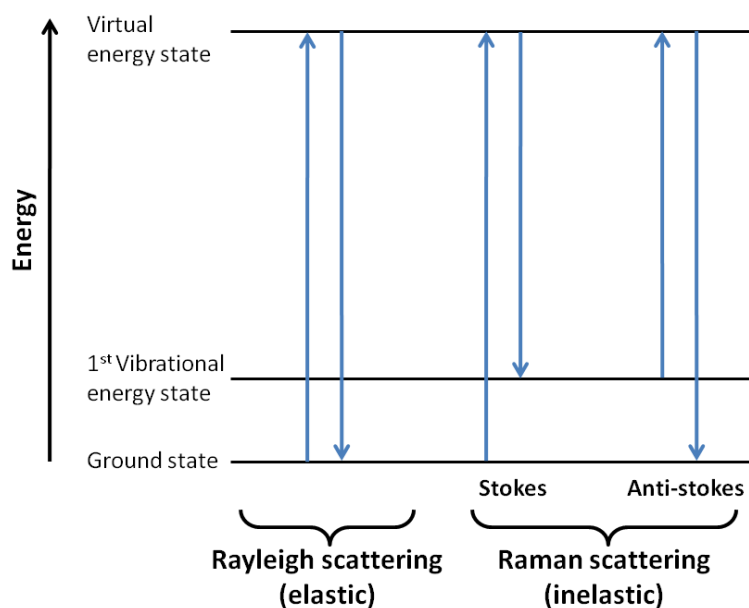


Figure 1.16. Diagram showing energetic transitions involved in Rayleigh and Raman scattering models. In the case of Rayleigh scattering an excitatory photon raises the system to a high (virtual) energy state and then a photon of equal energy (and therefore wavelength) is released as the system decays to the ground state. In Raman scattering the scattered photon is not equal in energy to the excitatory photon. This inelastic scattering can be divided into Stokes scattering, where the excited system decays to a vibrational energy higher than the ground state, and anti-Stokes scattering, where the initial excitation occurs from a higher vibrational state and the system decays back to the ground state. In the former case the wavelength of the scattered photon is longer than that of the excitatory photon and in the latter case it is shorter.

Resonance Raman spectroscopy is an advanced form of the technique which can overcome the problem of insensitivity by tuning the excitation wavelength to correspond with the molecular vibration[101]. This increases the amount of Raman scattering and can lead to an increase of signal to the order of 10^6 . This technique is particularly effective in systems with high-extinction electronic transitions wherein the excitation wavelength can be positioned within the corresponding band.

This is particularly applicable to haems with their Soret band, which can be excited by laser light between 400 - 450 nm. This allows the haem to be focussed on with no background signal from the rest of the enzyme. Consequently, studies of haemoproteins, including P450s, using resonance Raman spectroscopy are common[102-104]. The technique can reveal the spin and oxidation states of a haem centre along with details of conformational aspects of the outlying groups on the

porphyrin ring. These factors can have a dramatic effect on the overall electrochemistry of the enzyme.

Resonance Raman spectra are usually presented with relative signal intensity on the y-axis against the frequency scale wavenumber (measured in cm^{-1}) on the x-axis. Due to many, often overlapping, signals for a spectrum of a complex system such as a haem, multiple curve fitting is sometimes used to deconvolute and isolate individual bands. A typical spectrum of an oxidised P450 contains a strong band (ν_4) which is an oxidation state marker for Fe^{3+} haem. A pair of bands constituting the ν_3 feature are markers of spin-state in the haem iron d-electrons[105]. These bands can indicate in what proportion the high and low-spin species are present, and have been used to confirm a high-spin shift in substrate bound P450s[106]. Other signals of interest include the high frequency modes: ν_{10} , ν_{11} and $\nu_{\text{C=C}}$, which report on conformational and electronic aspects of the porphyrin vinyl and propionate groups[102]. These groups can have a large effect on the electronic properties of the haem and their conformation is often moderated by the position of side chains of amino acid residues in the P450 active site[13].

1.2.4. Electron Paramagnetic Resonance

Unpaired electrons are common in redox enzymes, present either during turnover (*e.g.* flavin semiquinone) or in the resting state (*e.g.* low-spin, Fe^{3+} haem iron). Electron paramagnetic resonance spectroscopy (EPR) is a powerful technique for characterising systems containing unpaired electrons (*i.e.* having a half-integral spin quantum number, $l = 1/2, 3/2, 5/2 \dots$). The principle is similar to that underlying nuclear magnetic resonance (NMR) spectroscopy, with resonance of electrons being measured rather than those of nuclei[107].

The species of interest containing unpaired electrons is placed in an applied magnetic field B_0 . The applied magnetic field causes the splitting of the electron spin moments into two populations: those aligned with the applied field ($+1/2$) and those opposed ($-1/2$) (Figure 1.17). The splitting energy is proportional to the applied field (Equation 1.7) and the constant of proportionality is the g-factor g_e (g-value).

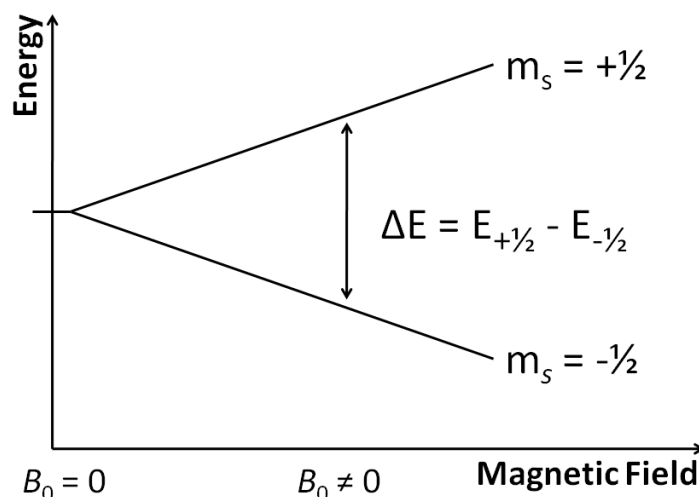


Figure 1.17. Representative diagram showing the splitting, due to the Zeeman effect, of the spin moments of electrons in an applied magnetic field B_0 .

$$\Delta E = g_e \mu_B B_0$$

Equation 1.7. Proportionality between the splitting energy of electrons in a magnetic field and the field strength (B_0). μ_B is the Bohr magneton and g_e is the g-factor.

Unlike NMR, where the magnetic field is fixed and radio frequency is used to probe the resonance, in EPR the sample is exposed to a constant frequency of microwave radiation while a varying magnetic field is applied. Typically 1st derivative spectra are shown and these are doubly integrated for analysis.

Several configurations of haem iron d-electrons are paramagnetic, and hence visible on an EPR spectrum (*e.g.* low-spin Fe^{3+}). EPR spectra can give clues about the coordination state of a haem centre, thus indicating the presence of molecules that are ligands or perturb the normal ligation state of the haem (*e.g.* substrate bound to a P450 inducing a change in spin state). Additionally, other redox cofactors with paramagnetic states and even radicals carried on amino acid residues of an enzyme can be investigated by EPR.

The technique can also be used to compare different P450 haem domains with one another in terms of the g-values of signals present. This can provide a qualitative insight into how similar the haem environments of different P450s are. In fusion P450s where flavin or iron sulfur cofactors are present, EPR is used to probe these centres as well.

1.2.5. Redox Potentiometry

Potentiometry is an electrochemical technique which looks at potentials, a measure of the difference in free energy between two oxidation states. The classical experimental set-up, the electrochemical cell, consists of two electrodes of different transition metals in the solid phase immersed in separate solutions of the same metals in a higher oxidation state. A salt bridge completes the circuit. The potential for electrons to move through the cell (E) can be measured using a simple voltmeter. The measured potential for a cell is the difference between the potentials for both halves, and so a potential measurement is always a comparative measure. In order to give a measure of the potential of an electrochemical half reaction at one electrode the standard hydrogen electrode (SHE) ($\text{H}_{2(\text{g})}/\text{H}^+_{(\text{aq})}$) is defined as having a potential of zero under standard conditions (298.15 K temperature, 1 atm pressure). The potential of a given electrode against the SHE is known as its standard electrode potential (E^0). Potentials are always quoted for the reductive half reaction even if the electrode is oxidising against the SHE, in this case the value of E^0 will be negative. In most experimental cases, setting up an SHE is quite impractical; therefore another reference electrode with a known and stable E^0 is used. Reference electrodes commonly used include the standard calomel electrode (SCE) ($E^0 = +242$ mV). In enzyme systems the propensity of redox centres to be reduced, by either simple chemical or biological reducing agents, can be measured by a similar technique[108]. This requires that a spectroscopic change, usually in the UV/Visible spectrum, upon reduction can be monitored. In a typical experiment a reducing agent is sequentially titrated against a solution of the enzyme of interest while the UV/visible spectrum and potential against a reference electrode are recorded after each addition. Redox enzymes are often susceptible to oxidation by atmospheric oxygen and thus potentials must be determined under anaerobic conditions[109]. Compounds, which mimic natural redox intermediates, known as mediators are often added to the experiment to shuttle electrons between non-physiological reducing agents and the enzyme which allows the system to equilibrate faster[108]. By plotting the values of absorbance at a given wavelength where the change occurs against the corresponding potentials, a reduction profile can be shown. This profile can be fitted, using a form of the Nernst equation, to derive a midpoint potential for a reduction event (Equation 1.9).

$$E = E^0 - \frac{RT}{nF} \ln K \left(K = \frac{[Red]}{[Ox]} \right)$$

Equation 1.8 The equilibrium Nernst equation where E is the observed potential, E^0 is the standard potential, R is the ideal gas constant, T is the absolute temperature, n is the number of electrons transferred, F is the Faraday constant and K is the equilibrium constant between reduced and oxidised species as shown.

$$A = \frac{(A_{min} + A_{max}) 10^{\left(\frac{E_{mid} - E}{RTF}\right)}}{1 + 10^{\left(\frac{E_{mid} - E}{RTF}\right)}}$$

Equation 1.9. A rearranged form of the Nernst equation for one electron reduction (expressed in terms of absorbance measurements) for use in fitting potential curves. A is the measured absorbance, A_{min} and A_{max} are the minimum and maximum absorbance values, E_{mid} is the midpoint potential, E is the observed potential and RTF is a constant combining the temperature, ideal gas and Faraday constants and a conversion factor from natural to base ten logarithms.

This is the electrode potential at which half of the population have become reduced. Values of the midpoint potential show how readily a redox centre in an enzyme is reduced under experimental conditions and by inference physiologically. By comparing potentials either between enzymes or for the same enzyme in different states (*e.g.* in the absence or presence of substrate), the electrochemical properties of an enzyme system can be explored.

It is known that the redox potential of P450 haem iron can be greatly increased in the presence of substrates (see section 1.1.6). Examples include P450cam (CYP101A1) and BM3 (CYP102A1) which both experience shifts of +100 – 150 mV in their haem iron redox potentials upon substrate binding [110, 111]. As discussed previously, this is a form of redox control which prevents non-productive haem reduction in the absence of substrate. Redox potentiometry allows quantification of these shifts in potential. Additionally, for fusion P450s, the potentials of other redox cofactors participating in the electron transport chain of the enzyme can be measured.

1.2.6. Differential Scanning Calorimetry

The fundamental principle of calorimetric experiments is to measure the thermal energy exchanged between a system and its surroundings. The most basic experiment that can be described as calorimetry involves burning a known amount of a substance and measuring the associated rise in temperature to determine the energy stored in the substance. This is essentially how the calorific values seen in the nutritional information of many food products are estimated.

Calorimetric studies of the type performed for this thesis are many times more sensitive than a simple burning experiment and measure changes in molar heat capacity (C_p) during the gradual unfolding of an enzyme.

For these experiments a sample of the enzyme concerned, in an appropriate adiabatic cell, is heated across a range of temperatures in small steps. For each step the amount of energy input required to achieve the temperature change is measured and from this C_p is calculated. From this value must be subtracted a value for a reference cell containing the same buffer but no protein, to give a differential heat capacity (ΔC_p).

As the protein sample is heated it unfolds. This may occur gradually, showing a smooth profile of heat capacity, or as a series of unfolding events which give rise to peaks in the heat capacity profile. In most cases, real proteins show a combination of these two phenomena, with some gradual unfolding to give a looser conformation and some large scale unfolding events. The temperature at which large scale events occur is known as a melting temperature (T_m). As well as revealing these physical properties of a protein, the technique can be used to probe the effects of different buffering conditions and ligands upon a protein by repeating the experiment in a range of buffers and observing changes in T_m values.

With respect to studies reported in this thesis, DSC is used to probe stability of P450s in different ligand-bound states, revealing substantial differences between different P450s and some stabilizing effects of bound ligands, as well as multiple unfolding events. DSC is also used for examining the relative stability of a fusion P450 and the constituent domains thereof. The technique has been used for the same purpose with BM3[112].

1.3. Enzyme Systems Studied

The techniques described in the previous section (1.2), along with others described in chapters 3-5, were used to characterise two P450 fusion enzymes. These two enzymes were CYP102A3 and CYP116B1. This section presents some background information regarding these two enzyme systems, with further details to be found in the general information sections of chapters 3 (CYP102A3) and 4 & 5 (CYP116B1).

1.3.1. CYP102A3

As described above (in section 1.1.9), the closest homologues of BM3 are two enzymes identified from the genome of *Bacillus subtilis*[113] which were given the systematic names CYP102A2 and CYP102A3.

Both are fatty acid hydroxylase enzymes sharing the same layout as BM3, with a P450 domain covalently linked to a CPR-like diflavin reductase. Amino acid sequence identity with BM3 is 64 % and 58 % for CYP102A2 and CYP102A3 respectively (ClustalW).

The two enzymes were first identified from the *Bacillus subtilis* genome in 2001 by Lee *et al.*[114]. The first purification from an *E. coli* expression system and the subsequent characterisation and kinetic studies were published in 2004 by Gustafsson *et al.* [79]. This work found a reduced affinity and turnover for saturated, straight-chain, fatty acid substrates and an increased affinity and turnover for branched-chain fatty acids for both enzymes compared to CYP102A1 (BM3). It was also found that these two enzymes both had altered regioselectivity compared to BM3 which favours myristic acid hydroxylation at the ω -1 position. The ratio of hydroxylated products for this substrate with BM3 at ω -1, -2 and -3 positions was 2.75 : 1.25 : 1, with CYP102A2 this ratio was 1.16 : 3.44 : 1 and for CYP102A3, 1 : 4.25 : 3.87[79]. This shows that (for myristic acid) CYP102A2 favours hydroxylation at the ω -2 position and CYP102A3 favours ω -2 and -3 positions almost equally. For three of the branched fatty acids (12 and 13-methylmyristic acids, 14-methylpentadecanoic acid), sigmoidal binding and kinetic dependencies were observed for both enzymes, suggesting cooperativity of binding for these substrates. The kinetics and regioselectivity of the reaction of the CYP102A1-3 enzymes with 12- and 13-methylmyristic acids was investigated[115]. This work found that the major products

of 12-methylmyristic acid for all enzymes were those hydroxylated at the ω -2 and ω -4 positions, with CYP102A1 and A2 favouring the former, and CYP102A3 favouring the latter. For the same substrate, minor products were in most cases shown for hydroxylation at every position between ω -1 and ω -6. For 13-methylmyristic acid all three enzymes strongly favoured hydroxylation at the ω -3 position with minor products hydroxylated at ω -2 and ω -4. It has been suggested that branched chain fatty acids are the natural substrate for BM3 and, possibly by extension, the whole CYP102 family[116]. There has also been work on CYP102A3, paralleling work on BM3, aimed at engineering the enzyme to improve and alter regioselectivity and change substrate specificity. One study produced a double mutant, based on equivalent residues in BM3 (F88V, S189Q), able to hydroxylate several organic compounds including *n*-octane and naphthalene[80]. Further work was able to focus the regioselectivity, of hydroxylation by CYP102A3, to different terminal and subterminal positions on a carbon chain[117].

Issues with loss of the haem cofactor shown for CYP102A2 seem to have focussed efforts (such as those discussed in the previous paragraph) on CYP102A3, whose haem is bound stably[79]. Following this reasoning CYP102A3 was chosen from the two for study in this thesis (more in section 3.1).

1.3.2. CYP116B1

The first of the CYP116 family of P450 enzymes now known as CYP116A1 was isolated from *Rhodococcus sp.* strain NI86/21 in 1995 by Nagy *et al.*[118]. It was found to be involved in the degradation of herbicides *S*-ethyl dipropylthiocarbamate (EPTC) and atrazine by this organism[119]. Three adjacent genes (CYP116A1, and two redox partners) are upregulated in the presence of EPTC[120] and CYP116A1 has detectable activity against both EPTC and atrazine. For both of these substrates hydroxylation leading to *N*-dealkylation is observed[118]. The redox partners were identified as a [2Fe-2S] ferredoxin (rhodocoxin)/ferredoxin (rhodocoxin) reductase type system.

Homology searches based on CYP116A1 identified a gene in the genome of the bacterium *Cupriavidus metallidurans* (formerly known as *Ralstonia metallidurans* which was reclassified from *Ralstonia eutropha* and *Alcaligenes eutrophus*)[121]. This organism is a soil bacterium of the order *Burkholderiales* and mainly studied for its ability to grow in millimolar concentrations of heavy metal ions including Ag⁺, AsO₂⁻, HAsO₄²⁻, Au⁺, Au³⁺, Bi³⁺, Cd²⁺, Co²⁺, Cu⁺, Cu²⁺, CrO₄²⁻, Hg²⁺, Ni²⁺, Pb²⁺, SeO₂³⁻, SeO₄²⁻, Sr²⁺, UO₂²⁺ and Zn²⁺[122-127].

The *CYP116B1* gene (designated Rmet_4932 in the genome project) is located the second of two megaplasmids that make up the *C. metallidurans* genome[128]. It encodes a 780 amino acid protein containing a P450 haem domain homologous to CYP116A1 and a domain resembling phthalate dioxygenase reductase. This reductase unit is composed of subdomains equivalent to the redox partners of CYP116A1 *i.e.* a [2Fe-2S] ferredoxin and a ferredoxin reductase. Phthalate dioxygenase reductase is a fairly well characterised enzyme with a crystal structure determined[129, 130] (Figure 1.18), kinetic data measured[131, 132], and investigation of the redox cofactors[133] published. It is the natural redox partner of phthalate dioxygenase, a non-haem iron-dependent, oxygenase enzyme that hydroxylates phthalate twice as part of a degradation pathway in bacterial species[134].

Hydroxylation of EPTC and related thiocarbamate pesticide vernolate by CYP116B1 has been demonstrated by HPLC analysis of reaction products[135].

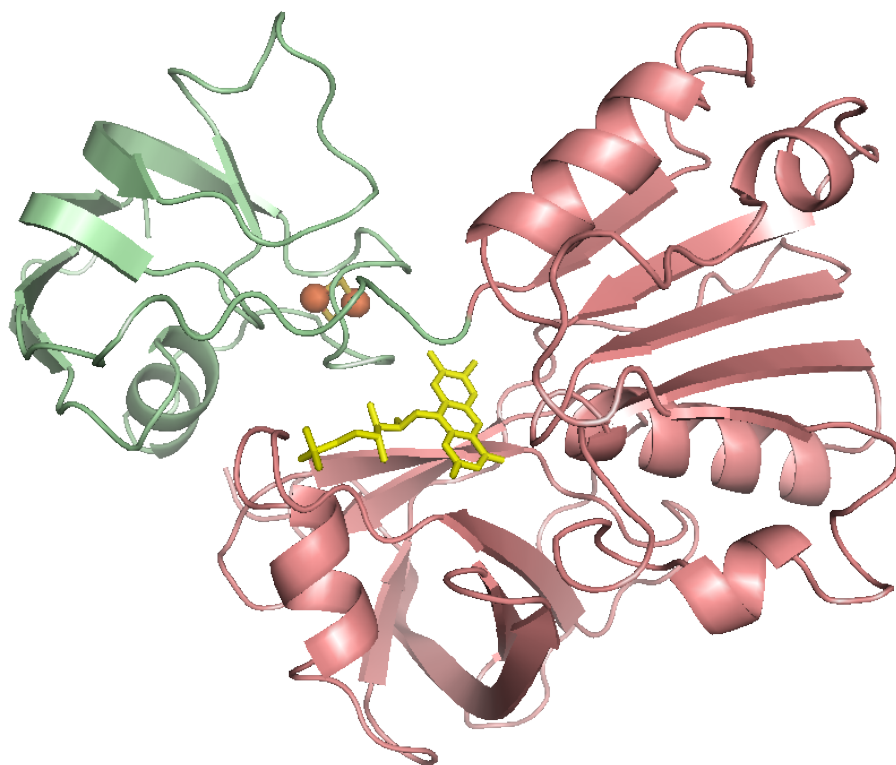


Figure 1.18 Image of phthalate dioxygenase reductase (PDOR) from the crystal structure (RCSB PDB accession number 2PIA). The ferredoxin-like subdomain is indicated in green and the ferredoxin-reductase-like subdomain in red. The redox cofactors bound to these subdomains are shown ([2Fe-2S] cluster, in orange and gold, and FMN, in yellow, respectively). The reductase domain of CYP116B1 is homologous to this enzyme.

1.4. Experimental Aims and Thesis Outline

The work on CYP102A3 is aimed at producing a crystal structure of the haem domain. A structure would allow an in-depth analysis of the structural basis of the differences in regioselectivity and substrate specificity of the CYP102A3 reaction compared to that of BM3. A structure of HDCYP102A3 might also inform mutagenesis work aimed at further understanding the structure function relationship of not only CYP102A3 but all enzymes of this subfamily. This structural work is to be supported by further biophysical characterisation of the HDCY102A3 protein using spectroscopic and potentiometric techniques.

CYP116B1 is of interest because the hydroxylation of thiocarbamate herbicides by the enzyme has been previously demonstrated. The work on this CYP116B1 is aimed at first purifying and characterising the enzyme and its isolated domains including

structurally by x-ray crystallography. The second, and arguably most important, aim is to investigate the hydroxylation of thiocarbamates by this enzyme using steady state kinetics to probe the rate of turnover and product determination techniques to confirm the identities of the products of the enzymatic reaction. This work could be a starting point towards understanding and developing this enzymatic activity for use in industrial applications.

The following chapters detail the experimental work carried out on the two P450 fusion enzymes described above. Chapter 2 describes the equipment and reagents used, and the experimental protocols followed in the course of the work. Chapters 3-5 are results chapters which present more detailed background information about the enzymes studied, the experimental data collected, and analysis and interpretation of these data. Chapter 3 details work on the haem domain of CYP102A3 with comparisons drawn with other P450 model enzymes. Chapter 4 begins with an examination of P450 fusion enzymes discovered from genome sequence data. These range from well characterised examples to putative novel fusion enzymes. This chapter continues by presenting the work concerning CYP116B1 and its constituent domains as proteins. This includes the optimisation of their purification, and a selection of biophysical techniques aimed at elucidating structural aspects. Chapter 5 presents the remainder of the work on CYP116B1 and its isolated domains, focussing on their enzymatic properties. This includes steady-state kinetic experiments, characterisation of reaction products. Chapter 6 contains an overall summary and discussion of all the results presented in the thesis, suggestions for further work on these enzymes, and the implications for the wider field of the study of the cytochromes P450.

Characterisation of Novel Cytochrome P450 Fusion Systems

Chapter 2

Materials and Methods

2. Materials and Methods

2.1. General Information

This chapter contains a comprehensive guide to the experimental protocols equipment and materials used to produce the work described in the other chapters of this thesis.

In the materials section, commonly used reagents and buffers are detailed along with some of the equipment used for various protocols.

The experimental methods section contains practical descriptions of the experiments reported in chapters 3-5, and how they were carried out.

2.2. Materials used

2.2.1. Microbiology Reagents

The antibiotics ampicillin and carbenicillin were supplied by Formedium (Hunstanton, UK). Rifampicin and kanamycin were supplied by Melford Laboratories Ltd (Ipswich, UK). Tetracycline was supplied by Sigma Aldrich (Poole, UK).

Antibiotic	Stock Concentration	Diluent	Typical Usage Concentration
Ampicillin	50 mg mL ⁻¹	dH ₂ O	50 µg mL ⁻¹
Carbenicillin	50 mg mL ⁻¹	dH ₂ O	50 µg mL ⁻¹
Kanamycin	30 mg mL ⁻¹	dH ₂ O	30 µg mL ⁻¹
Rifampicin	10 mg mL ⁻¹	0.17 M NaOH in 67 % MeOH	200 µg mL ⁻¹
Tetracycline	5 mg mL ⁻¹	50 % EtOH/dH ₂ O	12.5 µg mL ⁻¹

Table 2.1. Table of antibiotics used, details of the stock solutions prepared and the final concentrations at which they were typically used.

Isopropyl β-D-1-thiogalactopyranoside (IPTG) was made up as a 1 M stock in distilled, deionised water (dH₂O) and filter sterilised using a 0.2 µm syringe filter. Powdered pre-mixed growth media (LB, TB and 2xYT broths and LB agar) were made up as recommended by the supplier and then sterilised by autoclaving. IPTG and growth media were both supplied by Formedium.

2.2.2. Buffers

2.2.2.1. *Tris EDTA (TE) buffers*

50 mM Tris

1 mM EDTA

Ingredients were dissolved in dH₂O at 4 °C to around 80% of the final volume. The pH was then adjusted to the required value with concentrated hydrochloric acid. The buffer was then made up to the final volume with dH₂O.

2.2.2.2. *Phosphate Buffers*

50/500 mM potassium phosphate pH 6.5

1 M stocks of potassium dihydrogen phosphate (KH₂PO₄) and dipotassium hydrogen phosphate (K₂HPO₄) were made. These stock solutions were diluted to the required concentration of phosphate (50 or 500 mM) and the dibasic solution was titrated in to the monobasic solution until the required pH was obtained.

2.2.2.3. *Nickel Purification Buffers*

50 mM KPi pH 8.0

300 mM KCl

10% Glycerol (v/v)

Solid K₂HPO₄ and KH₂PO₄ were weighed out in proportions calculated using the Henderson–Hasselbach equation to give a final concentration of 50 mM phosphate and a pH of 8.0. These were dissolved with the KCl and made up to 90% of the final volume. Glycerol was then added to make the buffer up to the final volume.

For washing, elution and cleaning of nickel chromatography resins, various concentrations of imidazole were added to portions of this buffer. The solid imidazole was weighed out and dissolved directly into the buffer.

2.2.2.4. *Assay Buffers*

For HDCYP102A3, all characterisation was carried out in 100 mM KPi at pH 7.0 with 100 mM KCl. Kinetic assays of RDCYP116B1 were carried out in 50 mM KPi at pH 7.0

with 250 mM KCl. These buffers were produced using the same method as the nickel purification buffer but excluding the glycerol.

FLCYP116B1 kinetic assays, and product determination reactions, were carried out in 50 mM Tris buffer pH 8.0. This buffer was produced in the same way as TE buffers but excluding EDTA.

2.2.3. Protein Chromatography

2.2.3.1. Columns

Chromatography columns were used to purify protein samples for use in later experiments. Protein chromatography follows the fundamental principle of chromatography with a mobile phase moving through a stationary phase. In the case of protein chromatography the mobile phase is liquid, usually a buffer solution, and the stationary phase is a solid matrix or resin. The resin is composed of beads. This allows the resin to interact with components of the liquid phase while enabling it to permeate the resin bed. The liquid phase is commonly a buffer solution which is pushed through the resin either by gravity flow or using a peristaltic pump. Protein samples are loaded on the column in the liquid phase and can be separated and therefore purified based on their interaction with the solid phase. The composition of the liquid phase can also be used to modify the interaction between a species and the solid phase.

Column Name	Resin Name	Resin Supplier	Internal Diameter	Bed Volume	Column Supplier
P-seph	Phenyl Sepharose 6 Fast Flow	GE Healthcare	26 mm	64 mL	GE Healthcare
Q-seph	Q-Sepharose Fast Flow	GE Healthcare	26 mm	50 mL	GE Healthcare
DEAE	DEAE-Sepharose Fast Flow	GE Healthcare	50 mm	104 mL	GE Healthcare
Nickel	Ni-Superflow	Generon	30 mm	20 mL	Sarstedt
S-200	Sephacryl S-200	GE Healthcare	16 mm	104 mL	GE Healthcare

Table 2.2. Information about the chromatography columns used in the purification of protein samples.

Additionally small pre-poured, disposable gel filtration columns were employed as part of several techniques. These were EconoPac DG-10 (BioRad. Hemel Hempstead, UK).

2.2.3.2. *Linear Gradients*

Linear gradients used for the elution of columns were set up using a gradient mixer. This consists of two vessels connected by a tube with a valve. One vessel contains the running buffer and has another tube through which the gradient buffers are pumped onto the column. The other vessel contains the same volume of a buffer that will promote elution. This varies with the column being used. For ion-exchange columns (DEAE, Q-seph *etc.*) a buffer with high ionic strength can be used, and for hydrophobic interaction (P-seph) a buffer with low ionic strength was used. At the start of the gradient the stopcock between the two vessels is opened and rapid stirring of the first vessel is started. As the running buffer is pumped onto the column the elution buffer is pulled into the running buffer and mixed. This leads to a linear gradient between the running buffer and the elution buffer which has an endpoint approximately equal to the elution buffer. The length of the gradient can be altered by equally increasing the volumes of both buffers.

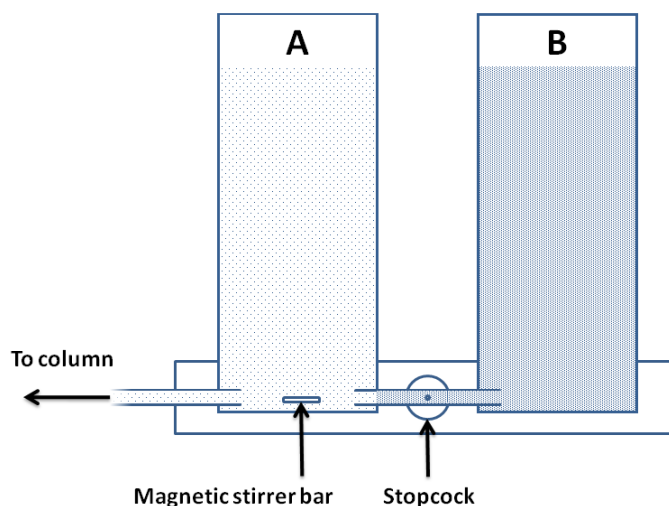


Figure 2.1. Schematic diagram of a gradient mixer. The vessel labelled A contains the running buffer for the column and the vessel labelled B contains elution buffer. The gradient is formed as buffer is pumped from vessel A onto the column and is replaced by buffer from vessel B as the vessels' volumes equilibrate. Rapid stirring in vessel A keeps the mix of buffers uniform throughout the vessel. The result of this process is a linear gradient between the two buffers.

2.3. Suppliers

2.3.1. CYP116B1 Substrate Candidates

The pesticides: atrazine, di-allate, EPTC, molinate, prosulfocarb, thiobencarb and vernolate were all from the PESTANAL range by Fluka Analytical supplied by Sigma Aldrich (Dorset, UK). Vanillic acid (99%) was from the Aldrich range also supplied by Sigma Aldrich.

2.3.2. Cell Strains

The *Escherichia coli* strains used in this work: NovaBlue (tolerates high plasmid copy number for DNA work) and BL21(DE3), HMS174(DE3), Origami (DE3) and Rosetta 2 (DE3) (T7 RNA polymerase lysogen expression strains) were all supplied by Novagen, a division of Merck Chemicals (Darmstadt, Germany).

2.3.3. Enzymes

Pfu turbo DNA polymerase, for PCR reactions, was supplied by Stratagene, a part of Agilent Technologies (Stockport, UK).

The following enzymes used in the work were supplied by New England BioLabs (Herts, UK):

Restriction enzymes AseI, BlnI, NdeI and NcoI for cloning and diagnostic digests.

T4 DNA ligase for ligation reactions.

2.4. Experimental Methods

2.4.1. Site-Directed Mutagenesis

Site directed mutagenesis was carried out using a QuikChange kit by Stratagene (a division of Agilent Technologies, Berkshire, UK). This required oligonucleotide primers designed to anneal to the part of the gene sequence to be modified and including the desired mutation. The primers were synthesised by Eurofins MWG Operon (London, UK). A polymerase chain reaction (PCR) was then set up to synthesise a mutated copy of the gene using an existing construct containing the wild-type gene. This was conducted according to the instructions supplied with the kit. After PCR amplification of the mutated vector, the template was destroyed by incubation with DpnI, a restriction enzyme that hydrolyses methylated DNA at its recognition site. Template DNA, produced by *E. coli*, is methylated whereas the artificial PCR product is not. Plasmids isolated from transformants were then confirmed by restriction digest (using a unique site built into the mutagenesis primers without affecting the coding sequence) or by DNA sequencing of the mutated region. Typically, the entire mutant plasmids were sequenced to ensure the presence of the desired mutation and the absence of adventitious mutations.

2.4.2. Ligation Reactions

Vector and gene insert DNA for ligation reactions were cut using the appropriate restriction endonuclease enzymes. The required vector and insert DNA fragments were then purified by gel extraction. This involved running them on a 0.8% agarose

gel containing ethidium bromide (3 μM), excising the appropriate bands (visualised using a UV transilluminator) and extracting the DNA using a gel extraction kit (QIAGEN, West Sussex, UK).

In the case of the CYP116B1 cloning, the insert DNA (the *CYP116B1* gene) was cut from a pGEM-T construct using *Asel* and *BlpI* restriction enzymes. The vector used was pET-15b (Novagen), which was prepared from a stock of a previous clone with the unrelated gene excised using *NdeI* and *BlpI* restriction enzymes.

Ligation reactions were set up in a 1.5 mL microfuge tube. Cut vector and gene insert DNA were added with T4 DNA ligase buffer concentrate (supplied with the enzyme) to sterile dH_2O for a final reaction volume of 40 μL . Lastly the T4 DNA ligase (200 units) enzyme was added to initiate the reaction. A mass of 100 ng of vector DNA was used for each reaction with various stoichiometric ratios of insert DNA between 2:1 and 1:5 across multiple reactions to maximise the chance of a successful ligation. Cut vector and gene insert DNA stock solutions were diluted so that additions were of around 1 μL in volume. Ligation reactions were left at room temperature for 16 hours.

A 1 μL aliquot of each reaction was then used to transform NovaBlue cells which were plated out onto LB agar plates containing ampicillin (100 $\mu\text{g mL}^{-1}$) which were incubated overnight at 37 °C. Colonies on the plates were miniprepmed to produce DNA stocks.

2.4.3. Transformation of Plasmid DNA into *E. coli* Cells

The required cell strain was streaked from a glycerol stock onto an LB agar plate and grown until single colonies formed (usually overnight growth was sufficient). A single colony from the streak plate was picked and inoculated into 5 mL of sterile LB growth medium in a universal tube (Sarstedt, Leicestershire, UK) and grown overnight. For each transformation a 1 mL aliquot of the overnight culture was dispensed into a 1.5 mL microfuge tube (Sarstedt) and pelleted by centrifugation in a benchtop microfuge at 13 000 rpm for 1 minute. The supernatant was poured away and the pellet resuspended in ice cold 50 mM CaCl_2 solution (0.5 mL). The cells were again pelleted by centrifugation as above and resuspended in ice cold 50 mM CaCl_2 solution (0.5 mL) a second time. The suspension was incubated on ice for 5 minutes before the addition of 1 μL of the transformant plasmid from a stock solution (typically containing ~ 1 ng

DNA). The transformation mixture was then incubated for 30 - 60 minutes at which point the cells were heat shocked at 45 °C for 30 seconds in a water bath. The heat shocked cells were then incubated on ice for a further 3 minutes and then 1 mL of SOC medium was added². The cells were once again pelleted and most of the supernatant poured away. With gentle pouring, 100-200 µL of residual supernatant remained in the tube and the pellet was resuspended in this residual fluid. Of the resuspended cells, 100 µL were spread onto an agar plate containing appropriate antibiotics. The plate was incubated at 37 °C until single colonies of transformed cells were seen (usually overnight).

2.4.4. Glycerol Stocks

Transformed cell lines were stored as glycerol stocks at -80 °C. These were made by inoculating 5 mL of sterile LB growth medium containing appropriate antibiotics with a single colony from a freshly grown transformation plate. This culture was grown at 37 °C with orbital shaking at 200 rpm until its optical density at 600 nm (OD_{600}) was 0.5 - 0.7, upon which a 200 µL aliquot was placed into an ice cold 1.5 mL microfuge tube. To these cells was added 800 µL of sterile 80% (v/v) glycerol and the glycerol stocks were then immediately frozen and stored at -80 °C. When the transformed cell line was required, a scrape from the corresponding glycerol stock was streaked onto an LB agar plate containing the appropriate antibiotics and the plate incubated overnight at 37 °C until single colonies were visible.

2.4.5. Amplification of Plasmid DNA

The plasmid of interest was transformed into the NovaBlue strain of *E. coli* (Novagen) as described above (2.1.1). A single colony of NovaBlue cells containing the plasmid was inoculated into 50 mL of sterile LB growth medium containing ampicillin and incubated overnight at 37 °C with orbital shaking at 200 rpm. The resultant cell culture was pelleted by centrifugation at ~4000 x g for 10 minutes in a benchtop centrifuge. The supernatant was poured off and plasmid DNA extracted and purified from the pellet using a Mini Prep kit (QIAGEN) according to the instructions supplied with the kit, scaled up for 50 mL culture volume. The resulting pure plasmid DNA could be confirmed by running on a 0.8% agarose gel containing ethidium bromide (3

² BL21(DE3), Origami(DE3) and Rosetta2(DE3) strains require incubation at 37 °C for an hour after the addition of SOC medium to ensure a high enough cell density and expression of antibiotic resistance for plating out.

μM) for 1 hour at 80 V. DNA in the gel was then visualised using a UV transilluminator.

2.4.6. Protein Expression Trials

Expression trials were conducted on a small scale with 50 mL of LB growth medium being inoculated with cells transformed with the appropriate expression vector plasmid construct. The cells were grown until their $\text{OD}_{600\text{S}}$ were in the region 0.6 - 0.8 at which point they were induced with IPTG (500 μM). Expression was continued for ~24 hours with 1 mL samples removed at intervals to monitor expression.

Their $\text{OD}_{600\text{S}}$ were recorded and the samples were then pelleted, by centrifugation in a benchtop microfuge at 13 000 rpm for 1 minute, and frozen. When all time point samples had been collected the cell pellets were lysed by resuspension in 50 μL of BugBuster solution (Novagen) and incubation therein for 30 minutes at room temperature. The lysed cells were then centrifuged in a benchtop centrifuge at 13 000 rpm for 10 minutes and the soluble fraction was pipetted off. The pelleted insoluble fraction was resuspended in 50 μL of TE buffer (50 mM Tris, 1mM EDTA, pH 7.2).

The soluble and insoluble fractions were examined by SDS polyacrylamide gel electrophoresis (SDS-PAGE) with the amounts loaded onto gels normalised according to the OD_{600} of the original samples. In this way the proportion of soluble protein expressed could be determined. These trials were conducted with variations in cell strain, expression temperature and IPTG concentration to optimise the protocol to be used in later large scale expression.

2.4.7. Expression and Purification of HDCYP102A3

Following a previous protocol for expression, an existing plasmid construct of pET21(+) containing the gene coding for the haem domain of CYP102A3 (HDCYP102A3) was transformed into BL21(DE3) cells (Novagen). This vector also contains a gene coding for resistance to the antibiotic ampicillin which was used for selection purposes. A single colony from the transformant plates was inoculated into 5 mL of sterile LB growth medium containing ampicillin (50 $\mu\text{g mL}^{-1}$) in a 30 mL tube. The 5 mL culture was grown for 6-8 hours and then 1 mL was inoculated into 200 mL of LB growth medium containing ampicillin (50 $\mu\text{g mL}^{-1}$) in a 1 L conical flask. The 200 mL culture was grown overnight at 37 °C with orbital shaking at 200 rpm and then 10 mL were inoculated into 500 mL aliquots of TB medium containing ampicillin

(50 $\mu\text{g mL}^{-1}$) in 2 L conical flasks, and these were incubated at 37 °C with orbital shaking at 200 rpm until the OD₆₀₀ of the culture was 0.6-0.8. HDCYP102A3 expression was then induced by the addition of 1 mM IPTG from a 1 M stock. Approximately 16 hours after induction the cells were harvested by centrifugation at 6 000 x g for 10 minutes. The pellet was then resuspended and washed in TE buffer to remove traces of growth medium, and then repelleted by centrifugation at 10 000 x g for 15 minutes. The cell pellet, which was visibly red most likely due to the presence of haem-containing HDCYP102A3, could then be frozen and stored at -20 °C until required. Typically a total of 12 L (24 flasks) of culture was grown at a time and the pellets combined.

The cell pellet was thawed and resuspended in TE buffer, and the protease inhibitors phenylmethylsulfonylfluoride and benzimidazole were added (1 mM each, final concentration). Keeping the resuspended pellet on ice, it was then lysed by sonication for a total time of 1 hour in cycles of 5 seconds on and 55 seconds off to avoid overheating (*i.e.* a total sonication time of 5 minutes). The cell lysate was centrifuged at 40 000 x g for 40 minutes to remove cell debris and the supernatant fluid containing HDCYP102A3 was decanted. This crude solution of HDCYP102A3 was treated with a 30% saturated solution of ammonium sulfate to precipitate some proteins and other components of the extract. Precipitated material was removed by centrifugation at 40 000 x g for 40 minutes. The ammonium sulfate concentration was then increased to 60% whereupon HDCYP102A3 was precipitated and could be separated by centrifugation as above. Treatments lasted 20 minutes. The 60% ammonium sulfate pellet was resuspended in TE buffer to an ammonium sulfate concentration of approximately 1.5 M (usually around 250 mL). The HDCYP102A3 solution was then bound to a phenyl Sepharose (P-seph) column pre-equilibrated with TE buffer containing 1.5 M ammonium sulfate and washed with 2-3 column volumes of the same. The protein was then eluted by application of a linear gradient of ammonium sulfate in the concentration range 1.5 M – 0 M and the eluate was collected into 2 mL fractions. Fractions were selected visually, with the most intensely coloured ones pooled and extensively dialysed against TE buffer to remove residual ammonium sulfate. The dialysed HDCYP102A3 was concentrated to approximately 50 mL and then loaded onto a quaternary-amine Sepharose (Q-seph) column, of 30 mL bed volume, pre-equilibrated with TE buffer. The column was then washed with 2-3 column volumes of TE buffer and the HDCYP102A3 was eluted by

application of a linear gradient of potassium chloride in TE buffer (0-500 mM) and the eluate was collected into 2 mL fractions. The fractions were analysed using a UV/Vis spectrophotometer to measure the ratio of absorbance due to haem (~420 nm) and protein aromatic amino acids (~280 nm) as a rough measure of the purity of HDCYP102A3. Those fractions with a $A_{420}:A_{280}$ greater than 1:1 were pooled and dialysed against TE buffer to remove the potassium chloride.

For crystallographic trials the HDCYP102A3 was further purified using a Sephadex S-200 size exclusion column (GE Healthcare, Buckinghamshire, UK). The column was equilibrated with two column volumes of 10 mM Tris pH 7.0. The HDCYP102A3 solution was concentrated using spin concentrators (Vivaspin 20) and a sample (200 μ L – 1 mL) was loaded directly on the resin bed using a glass Pasteur pipette. The column was then run with 10 mM Tris pH 7.0 and the eluate was collected into 1 mL fractions. Those fractions with a ratio $A_{420}:A_{280}$ greater than 1.4 were pooled and reconcentrated. This ultra pure stock of HDCYP102A3 was stored on ice until required for crystallisation trials.

2.4.8. Expression and Purification of FLCYP116B1 and HDCYP116B1

The pFLCYP116B1-15b (pHDCYP116B1-15b) construct was transformed into HMS174 (DE3) cells following the method described in section 2.4.3. A single colony from the transformation plate was used to inoculate 5 mL of sterile LB medium containing 50 μ g mL⁻¹ each of rifampicin and carbenicillin. The culture was grown for ~8 hours at 37 °C with orbital shaking at 200 rpm. A 100 μ L aliquot of the 5 mL culture was used to transfect 300 mL of sterile LB medium, in a 1 L conical flask, containing 50 μ g mL⁻¹ rifampicin and carbenicillin. The culture was grown overnight at 37 °C with orbital shaking at 200 rpm. 10 mL aliquots of the overnight culture were used to inoculate 600 mL cultures of sterile 2xYT, in 2 L flasks, containing 50 μ g mL⁻¹ ampicillin. Typically, 24 flasks were used for a total culture volume of 14.4 L. The culture was grown at 37 °C, with orbital shaking at 200 rpm, until the OD₆₀₀ was between 0.6 and 0.8. When the culture had reached this OD threshold, the incubation temperature was reduced to 25 °C and 300 μ L of 1 M IPTG was added to each flask to induce expression. Post-induction, the culture was incubated for at least 20 hours before cells were harvested in the same way as described for HDCYP102A3, above. Cell pellets were stored at -20 °C until required.

Thawed cell pellets (kept in an ice bath at all times) were diluted in nickel buffer to 250 mL and protease inhibitors, PMSF and benzamidine, were added. The cells were then lysed by sonication for 1 hour in cycles of 5 seconds on and 55 seconds off to avoid overheating (*i.e.* a total sonication time of 5 minutes). Cell debris was removed by centrifugation at 40 000 x g for 40 minutes. This operation and all further actions performed on the protein were carried out at 4 °C unless specified otherwise. The cell supernatant was loaded directly onto a 6xHis tag affinity column (as detailed in section 2.2.3.1). The column was washed five times with 20 mL nickel buffer before the protein was eluted with ~40 mL of 100 mM imidazole in nickel buffer. Eluent containing the FLCYP116B1 (HDCYP116B1) protein was collected in 5 mL fractions. The fractions were examined spectrally and by SDS-PAGE and those containing the most protein were pooled. The pooled fractions were concentrated using spin concentrators to a 2-5 mL volume, and dialysed against 5 L of nickel buffer twice to remove imidazole. The protein was then further purified using a size-exclusion column (S-200, detailed in section 2.2.3.1) run in nickel buffer. The eluent was collected as 1 mL fractions and those containing the peak elution (as determined spectrally) were pooled. The protein was then concentrated to 1-2 mL volume and dialysed into 50% glycerol in nickel buffer before being stored at -20 °C until required.

2.4.9. Expression and Purification of RDCYP116B1

The pRDCYP116B1-11a construct was transformed into HMS174 (DE3) cells as described in section 2.4.3. The culture was scaled-up and expression induced as described for FLCYP116B1 and HDCYP116B1 above. Cells were harvested 24 hours after induction of expression in the same way as described above and cell pellets were stored at -20 °C until required.

Pellets were thawed, and diluted in ice cold TE buffer (pH 7.2 throughout) on ice and protease inhibitors, PMSF and benzamidine, were added. The resuspended pellet was then lysed by sonication and had cell debris removed in the same way as described above for FLCYP116B1 and HDCYP116B1. The cell supernatant was treated with 40% saturated ammonium sulfate solution added slowly with stirring at 4 °C. The precipitated material was removed by centrifugation at 40 000 x g for 40 minutes and the pellets were discarded while the supernatant was retained. The ammonium

sulfate supernatant was loaded onto a phenyl Sepharose column pre-equilibrated with 1 M $(\text{NH}_4)_2\text{SO}_4$ in TE buffer. The column was then washed with two column volumes of 1 M $(\text{NH}_4)_2\text{SO}_4$ in TE buffer and two column volumes of 0.75 M $(\text{NH}_4)_2\text{SO}_4$ in TE buffer. The elution of RDCYP116B1 was then effected by application of a linear gradient of 0.75 M – 0 M $(\text{NH}_4)_2\text{SO}_4$ in TE buffer. Complete elution was often only achieved by washing with TE buffer after the linear gradient had completed. The eluent was collected as 5 mL fractions. The most visibly coloured fractions were pooled and dialysed against 5 L TE buffer twice to remove residual $(\text{NH}_4)_2\text{SO}_4$. The RDCYP116B1 protein was then concentrated, using centrifugal concentrators (molecular weight cut off: 10 kDa), to ~5-10 mL and loaded onto a Q-seph column pre-equilibrated in TE buffer. The column was washed with 2-3 column volumes of TE buffer and elution of RDCYP116B1 was effected by application of a linear gradient of 0-500 mM KCl in TE buffer. The eluent was collected in 2 mL fractions. The fractions were examined spectrally and those containing the most RDCYP116B1 were pooled, dialysed and concentrated as previously. The RDCYP116B1 protein was further purified using a ResourceQ column on an AKTA FPLC system (both GE Healthcare, Buckinghamshire, UK) using a similar protocol to that used for the previous Q-seph column. The purest fractions, identified by SDS-PAGE, were pooled and dialysed as above and then concentrated to a volume < 1 mL using centrifugal concentrators. The RDCYP116B1 protein was then dialysed against 50% glycerol in TE buffer and stored at -20 °C until required.

2.4.10. Multi-Angle Light Scattering (MALS)

The protein of interest was diluted to 5 μM concentration in its respective assay buffer and a sample was run on a size exclusion chromatography (SEC) column (S-200) eluting through a flow cell in the MALS instrumentation. This consisted of a DAWN-EOS MALS spectrometer (Wyatt Technology corp. Santa Barbara CA, USA) detector and a quasi-elastic light scattering (QELS) detector (Wyatt). MALS data were recorded, processed and fitted (using the Zimm fitting method[136]) using the software associated with the instrument (Astra, also by Wyatt). The use of this SEC-MALS type of experiment allowed the identification of components of the protein solution by retention time on the column as well as measurements of hydrodynamic radius and molecular weight from the MALS equipment.

MALS experiments were conducted by Mrs. Marjorie Howard at the Biomolecular Interactions Facility in the Faculty of Life Sciences, University of Manchester.

2.4.11. Crystallisation Trials

Initial crystallisation trials were conducted on the nanolitre scale using a mosquito pipetting robot (Molecular Dimensions, Newmarket, UK). Protein samples were tested against crystallisation screens composed of common crystallisation conditions (QIAGEN, Crawley, UK). The trials were conducted in 96-well crystallography trays which have 96 identical units set out in an 8x6 configuration. Each unit is comprised of three wells, a deep well to hold the crystallisation screen solution and two smaller, concave wells which hold the crystallisation drops (Figure 2.2). Crystallisation screens in the 96-well block format (QIAGEN) were pipetted by hand into the deep well on the 96-well crystallography trays. The Mosquito robot then pipetted 100-200 nL of 10-15 mg/mL protein solution into the top wells and then added an equal amount of the screen solution to create a 1:1 mixture of the two. The whole tray was then sealed with clear tape (Duck Tape, Avon OH, USA) and left to equilibrate at 4 °C. After a minimum of 24 hours equilibration, the wells were examined for signs of crystallisation using a stereoscopic light microscope (Nikon).

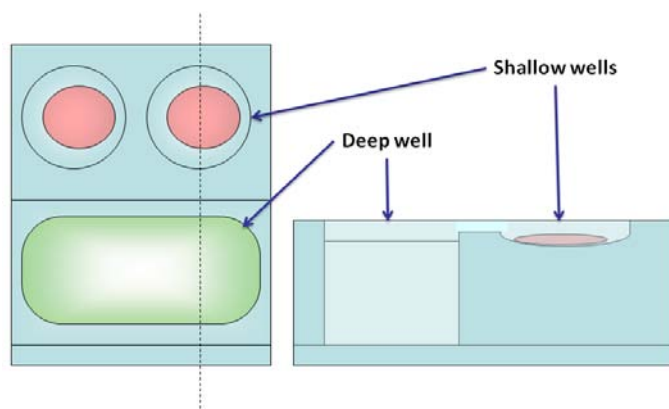


Figure 2.2. Representative diagram of a well unit from a 96-well crystallography tray seen from above (left) and as a side-on cross-section (right). At the bottom is the deep well containing the screening solution and above are the two wells containing the crystallisation trial drops. The deep well and the drops are open to each other to allow equilibration.

Secondary trials were set up based on any hits from the initial trials, on a larger microlitre scale, to optimise the crystallisation conditions.

2.4.12. Binding Titrations

Ligand binding to P450 enzymes was investigated by carrying out optical binding titrations. Titrations were carried out in a quartz cuvette into which the enzyme was diluted to 5-10 μM concentration in an appropriate buffer to a final volume of 1 mL (typically Tris or KPi buffers were used). The ligand under investigation was then titrated in from a concentrated stock in 0.1 - 1 μL steps using a gas-tight syringe of 1 or 10 μL capacity (Hamilton). The maximum volume added to an assay was 10 μL .

The UV/Visible spectrum (250-800 nm) of the assay was recorded initially and then after each addition of ligand. In this way sequential changes in absorbance associated with the binding of ligand could be observed. Spectral changes induced after each ligand addition were determined by subtraction of the absorption spectrum for the ligand free enzyme from that at the relevant titration point. Maximal absorption changes induced were defined as those at the peak and trough of the resultant difference spectrum. The total difference in absorption was computed by subtraction of the trough absorption from the peak absorption at these wavelengths. The same process was done for each point in the titration. These absorbance difference values were then plotted against the relevant ligand concentration and fitted (using either a standard hyperbola, Hill equation or the Morrison equation defined above, Equation 1.4, Equation 1.6) to generate a dissociation constant (K_d) value for the ligand (or apparent dissociation constant (K_H) in the case of the Hill equation).

2.4.13. Redox Potentiometry

This technique was used to measure the midpoint potentials of the redox centres present in the enzymes studied. The experiments were carried out in a 5 mL beaker, with a water jacket connected to a water bath for temperature control. A Cary 50Bio UV/Visible spectrophotometer (Varian, Palo Alto CA, USA) with a fibre optic probe was used to monitor changes in absorbance and a potentiometer (Mettler Toledo, Columbus OH, USA) with a calomel reference electrode attached was used to measure the potential as the experiment proceeded. All equipment was contained in a glovebox (Belle Technology, Weymouth, UK) under a nitrogen atmosphere to ensure anaerobic conditions. This excluded the effect of atmospheric oxidation from the experiment and in P450 enzymes prevented oxygen dependent turnover and reoxidation of ferrous haem iron. A sample (200-500 μL) of the protein of interest was passed through a small gel filtration column in anaerobic buffer to help remove

traces of dissolved oxygen. A volume of the protein sample collected from the column sufficient to produce a final concentration of around 5 μM was diluted to a volume 4.99 mL in buffer. The mediator compounds HNQ (hydroxynaphthoquinone), PMS (phenazine methosulfate), methyl viologen and benzyl viologen were added as shown in Table 2.3 to bring the total volume to 5 mL.

Mediator	Final concentration	Stock Solution
Hydroxynaphthoquinone (HNQ)*	0.3 μM	7 mM
Phenazine methosulfate (PMS)*	2 μM	5 mM
Benzyl viologen (BV)	1 μM	2.5 mM
Methyl viologen (MV)	0.3 μM	1.5 mM

Table 2.3. List of mediator compounds used in redox potentiometry experiments. Stock solutions of mediators marked (*) were prepared freshly on the day of use and stored in foil covered tubes to prevent photodegradation. All other mediators are stable and could be prepared beforehand.

The UV/Visible spectrum (800 – 250 nm) was recorded before and after the addition of mediators and the potentiometer was switched on. A solution of sodium dithionite, a powerful reducing agent, was then added in small aliquots to facilitate progressive reduction of the sample enzyme. After each addition the system was allowed to equilibrate until a stable reading was seen on the potentiometer. The UV/Visible spectrum was then recorded along with the corresponding reading of the potential. Repeated additions of dithionite were continued until no change in potential was seen after three or so further additions, usually at a maximum of around -650 mV vs. the Calomel electrode. The enzyme was then reoxidised to the starting potential by 3-6 additions of a solution of potassium ferricyanide ($\text{K}_3[\text{FeCN}_6]$) to show reversibility and absence of hysteretical behaviour. Absorbance data (usually at the Soret peak or at a wavelength of maximal absorbance change between oxidised and reduced enzyme samples) were then plotted against the applied potential (corrected versus the standard hydrogen electrode, SHE, by addition of +244 mV to the value relative to the Calomel electrode) and fitting was done using the Nernst equation to derive the midpoint reduction potential for the $\text{Fe}^{3+}/\text{Fe}^{2+}$ haem iron couple.

2.4.14. Reduced-CO Complex Formation

A sample of the purified protein (approximately 5 μM) was placed in a quartz cuvette and its UV/Visible spectrum (250-800 nm) was recorded. The protein was then

reduced by the addition of a few grains of solid sodium dithionite ($\text{Na}_2\text{S}_2\text{O}_4$) and the spectrum recorded again. The solution was then pulled carefully into a plastic syringe containing gaseous carbon monoxide and gently agitated. The solution was then dispensed from the syringe back into the cuvette with care being taken not to introduce bubbles. A layer of carbon monoxide from the syringe was introduced into the top of the cuvette and it was sealed with a rubber suba seal. Spectra were then recorded at 1 minute intervals so that the formation of the CO complex could be observed.

2.4.15. Preparation of FLCYP116B1 for kinetic assays

As loss of FMN cofactor had been noted during purification of FLCYP116B1, as a yellow colour in dialysis and column flow-through, the protein was incubated with FMN prior to kinetic assays. FLCYP116B1 was diluted to $\sim 40 \mu\text{M}$ in 1 mL of assay buffer which was combined with 1 mL of saturated FMN solution and incubated on ice overnight. The free FMN was removed immediately before commencing kinetic assays by passing the mixture through a small gel filtration column (see section 2.2.3.1) running in assay buffer. FLCYP116B1 passed quickly through the column and was typically collected in a volume of 0.5-1 mL at a concentration of 30-50 μM . The smaller free FMN moved more slowly through the column and was completely washed off with ~ 100 mL assay buffer.

2.4.16. Steady State Assays

Steady state enzyme kinetic assays were carried out using a Cary 50Bio UV/Visible spectrophotometer (Varian, Palo Alto CA, USA) to monitor the consumption (oxidation) of NAD(P)H coenzyme by a decrease in absorbance at a wavelength of 340 nm ($\Delta\epsilon_{340} = 6.210 \text{ mM}^{-1} \text{ cm}^{-1}$). From this absorbance change an initial rate of the enzyme reaction could be inferred after taking into account any background rate of NAD(P)H oxidation in the absence of any exogenous substrate. Disposable cuvettes of 1.5 mL volume and 1 cm pathlength suitable for UV work were used to contain and observe assays. All assays could be initiated by the addition of NAD(P)H coenzyme so that the enzyme and oxidising substrate were diluted at the appropriate concentrations in an assay buffer. Final assay volume was set at 1 mL.

After addition of the NAD(P)H coenzyme, the cuvette was covered with parafilm and inverted three times to mix assay components thoroughly. The cuvette was then placed into the spectrophotometer and the recording of absorbance data initiated.

These actions were performed rapidly but accurately to ensure that the reaction could be observed from as close to the initiation point as possible.

In some cases (where required) steady state assays were carried out using stopped flow apparatus (TgK Scientific, Bradford-on-Avon, UK). This presented several advantages over a simple UV/Visible spectrophotometer. Firstly, the volume of each assay was less, allowing preservation of enzyme stocks. Finally, in some cases standard steady state assays using a simple spectrophotometer produced erroneous results. This appeared in the form of traces with extreme noise and this was thought to be due to mixing effects. The use of stopped flow apparatus for these assays eliminated the problems and produced clean reproducible traces.

Stopped flow apparatus essentially consists of three syringes, two drive syringes and a stop syringe (Figure 2.3). The drive syringes are loaded with separate solutions of the enzyme and substrates. When the experiment is initiated the drive syringes are driven by compressed nitrogen gas pushing an equal amount of the two solutions through a rapid mixer and into the stop syringe through a cell with an absorbance detector. When the stop syringe fills the flow is stopped and a data trigger activated which begins the recording of absorbance data.

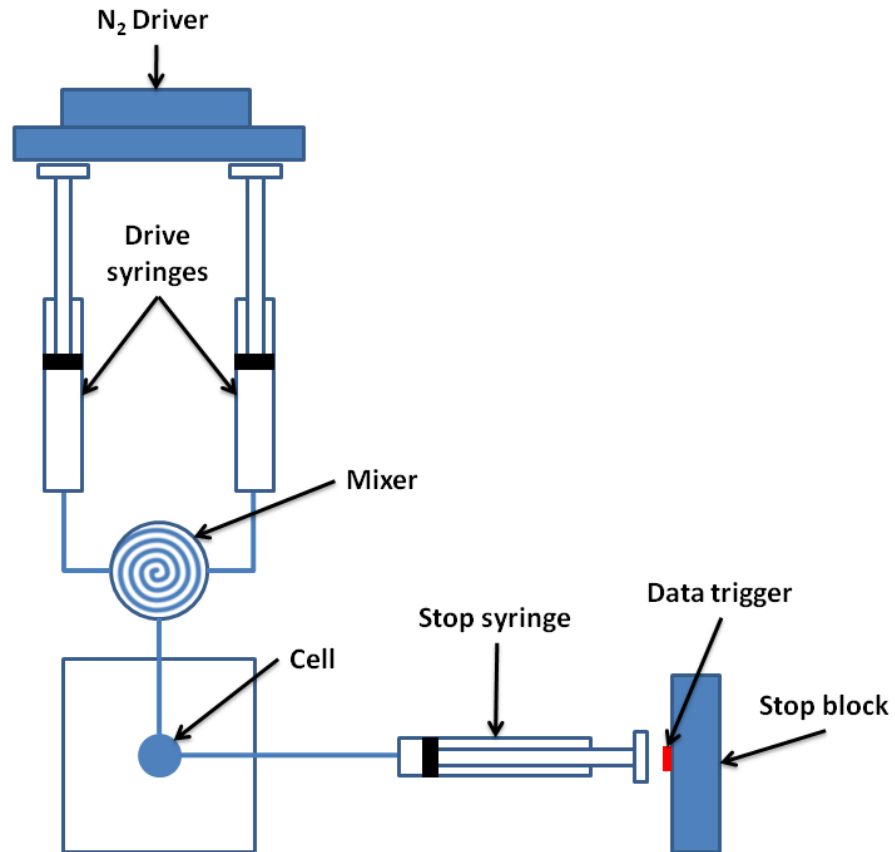


Figure 2.3. Representative diagram of a basic stopped flow apparatus. The two drive syringes contain separate solutions of enzyme and substrates at double assay concentrations. The drive syringes are driven simultaneously by compressed nitrogen gas to deliver equal volumes of both solutions. The flow is forced through a rapid mixer and into a cell where the reaction can be observed (usually by absorbance or fluorescence spectroscopy) and into the stop syringe. When the stop syringe fills it hits the stop block and activates the data trigger which begins the recording of data.

2.4.17. Electron Paramagnetic Resonance Spectroscopy

2.4.17.1. Preparation of EPR Samples

The protein stock of interest was diluted to a concentration of 200 μM in a volume of 200-400 μL of the preferred buffer (HDCYP102A3 – TE buffer) (FL and HDCYP116B1 – 50 mM KPi) along with any ligands required. For reduced spectra ~0.2 - 0.4 mg solid sodium dithionite was added to the sample. The prepared sample was then transferred to an EPR tube using a 1 mL syringe and canula. The sample was then carefully frozen in the tube using liquid nitrogen and then stored in the same until spectra were recorded.

2.4.17.2. EPR of HDCYP102A3

EPR spectroscopy of HDCYP102A3 was carried out by Dr. Myles Cheesman at the Chemistry department of the University of East Anglia, Norwich. Spectra were recorded on an EPR spectrometer comprising a ER-200D electromagnet and microwave bridge interfaced to an EMX control system (Bruker, Coventry, UK) and fitted with a ESR-9 liquid helium flow cryostat (Oxford Instruments, Oxfordshire, UK) and a ER4116DM type dual-mode X-band cavity (Bruker). Spectra were recorded at 10 K with 2 mW microwave power and a modulation amplitude of 10 G. Derivative spectra were generated from raw microwave absorption data by the software associated with the spectrometer (Bruker).

2.4.17.3. EPR of CYP116B1 and its Haem Domain

EPR spectroscopy of FLCYP116B1 and HDCYP116B1 proteins was carried out by Dr. Stephen Rigby of the Faculty of Life Sciences at the University of Manchester. Spectra were recorded using a ELEXSYS E500 spectrometer (Bruker) in conjunction with an ESR900 cryostat and ITC503 temperature controller (Oxford Instruments). Spectra were recorded at 10 K with 2.08 mW microwave power and a modulation amplitude of 10 G. Derivative spectra were generated from raw microwave absorption data by the software associated with the spectrometer (Bruker).

2.4.18. Differential Scanning Calorimetry

The DSC apparatus consists of the calorimeter and a computer that controls the experiment. A schematic diagram of the apparatus and some explanation of its functioning are presented in Figure 2.4 and the accompanying legend.

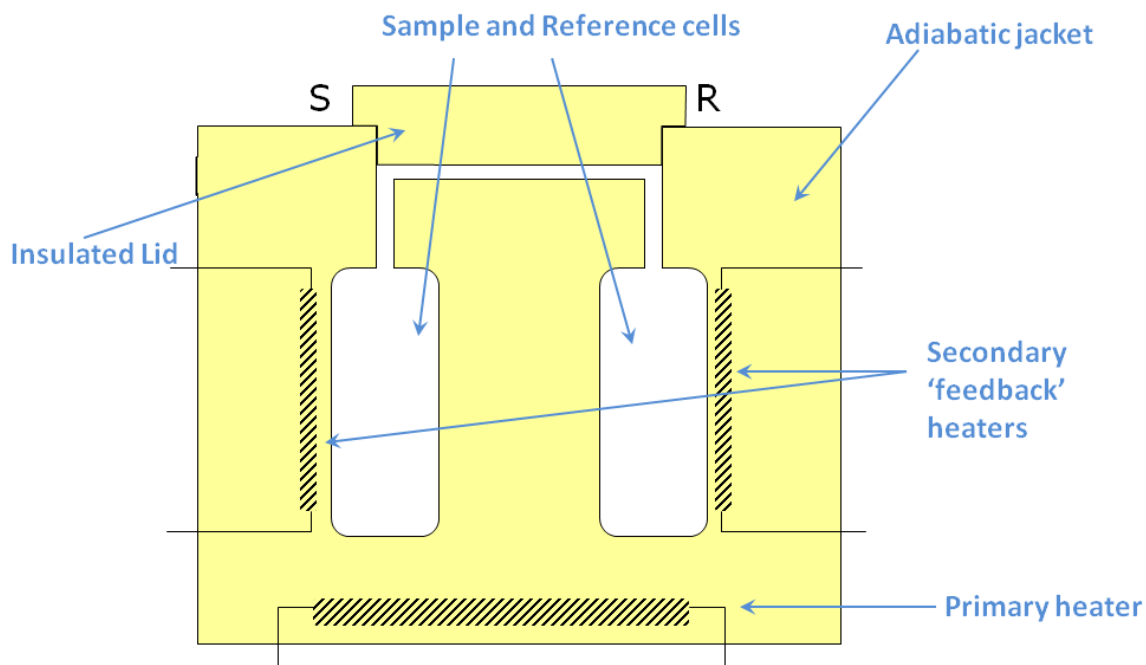


Figure 2.4. Schematic, labelled diagram of a differential scanning calorimeter such as that used for the experiments in this thesis. With sample and reference solutions in the appropriate cells, the insulated lid can be tightened to exert a positive pressure on the cells to prevent boiling during the experiment. The primary heater increases the temperature of the cells across the experimentally defined range. The ‘feedback’ heaters are used to determine the difference in the amount of energy required to cause incremental increases in temperature for each cell, and thereby, the differential heat capacity between them (dC_p).

Protein samples were prepared for DSC by dialysis into an appropriate buffer: 100 mM KCl in TE buffer, pH 7.2 for HDCYP102A3 and 50 mM Tris, pH 8.0 for CYP116B1 proteins. A sample of the dialysate was retained and filter sterilised for use as the reference. The instrument used was a VP-DSC by Microcal (a division of GE Healthcare, Buckinghamshire, UK), controlled using the supplied software.

Prior to use, protein and reference solutions for use in the DSC experiments were extensively degassed using a ThermoVac degassing unit (Microcal). The wells of the instrument were filled with 0.8 mL using a 1 mL syringe with a 10 cm narrow bore needle. At all times care was taken to avoid any aeration of the samples while filling the instrument. During the calorimetric experiments the instrument was sealed and pressurised to ~30 psi to prevent boiling of the samples. Each experimental run

involved ramping the temperature from 20 °C to 95 °C at a rate of 1 °C min⁻¹ (hence over a period of ~75 minutes).

First, a baseline was determined by filling both the sample and reference wells of the calorimeter with the reference buffer (the dialysate mentioned above). The two samples of buffer were subjected to consecutive runs until a consistent profile for the reference-reference measurement was achieved (typically after 4 or 5 runs). The experiment was then paused, and the sample cell was emptied and refilled with the dialysed protein sample. The experiment was then resumed and the unfolding profile of the protein of interest was recorded in one run. The raw profile was baseline corrected by subtracting the reference-reference profile and applying a further baseline correction. Curve fitting was then used to isolate the individual unfolding events making up the profile.

2.4.19. Resonance Raman Spectroscopy

For Resonance Raman spectroscopy, the P450 protein of interest was diluted to a suitable concentration (typically ~200 µM) and required ligands were added at a sufficiently high final concentration to enable near-saturation of the P450.

Excitation was done using 15-milliwatt radiation from an Innova 300 krypton ion laser (Coherent inc. Santa Clara CA, USA), at a wavelength of 406.7 nm, close to the low-spin haem Soret maxima of the P450 enzymes studied. Spectra were acquired using a micro-Raman system 1000 spectrophotometer (Renishaw, Gloucestershire, UK).

Spectra were baseline corrected and peak-fitted using GRAMS Spectral Notebase software (ThermoFisher Scientific, Hertfordshire, UK).

2.4.20. Product Determination from Prospective Substrate Oxidation by CYP116B1

The product determination experiments consisted of three distinct phases: The CYP116B1 enzyme reaction, product purification and analysis by HPLC-MS. These steps are described in the following subsections.

2.4.20.1. CYP116B1 Reaction

Reactions were set up in 30 mL brown glass vials (Supleco, a division of Sigma Aldrich) which had been washed with methanol, dichloromethane and dH₂O. Each

reaction contained one of the CYP116B1 substrate candidates at a concentration of 0.33 mM, added from a 100 mM ethanol stock. The reactions also contained 1 mM NADPH and 6 μ M CYP116B1 enzyme, both added from stock solutions in assay buffer. An NADPH regeneration system was also added, consisting of 8.25 activity units of glucose-6-phosphate dehydrogenase and 20 mM glucose-6-phosphate. The reactions were made to a final volume of 10 mL in assay buffer. Reactions were allowed to proceed, with slow agitation, for 16 hours, at which point the reaction was stopped by acidification with the addition of 0.5 mL 1.0M HCl. Any products in the reaction mixture were then purified as described below.

2.4.20.2. Product Purification

Any protein aggregate debris, produced by the acidification step, was removed by centrifugation at 4500 x g for 40 minutes. The product compounds were then purified from the clarified reaction mixture using Strata-X solid phase extraction columns (Phenomenex inc. Macclesfield, UK). This product employs a reversed-phase silica type resin to separate organic compounds by hydrophobic interactions. The columns were used in accordance with the accompanying directions and produced an elution of the product compounds in 1 mL of a methanol-acetonitrile mixture, which was transferred to sealed HPLC vials.

2.4.20.3. High Performance Liquid Chromatography-Mass Spectrometry

The products from the CYP116B1 reactions with substrate candidates were analysed by HPLC-MS. The analyses were performed using an Agilent 1100 series HPLC (Agilent Technologies, Berkshire, UK) linked to a LC/MSD SL instrument (Agilent), using a Exclipse XDB-C₁₈ HPLC column (Agilent). The experimental design was managed using the software supplied with the instrument (LC/MSD ChemStation by Agilent). A 10 μ L sample of the extracted products was eluted through the column in a gradient from 66-100% acetonitrile in water (HPLC grade solvents by Romil Ltd. Cambridgeshire, UK), both containing 0.1 % formic acid, at a flow rate of 1 mL min⁻¹ and a temperature of 30 °C. Eluate from the column flowed through to the mass spectrometer where the mass spectrum was recorded. The eluate was monitored by its absorbance at 254 nm and 220 nm, and the signal from the mass spectrometer detector for 68 minutes. This allowed the retention times for various components of the reaction mixture to be determined, and the mass spectra for these components to

be displayed. The HPLC-MS work was carried out with the kind assistance of Dr Bettina Nestl of the Department of Chemistry at The University of Manchester.

Characterisation of Novel Cytochrome P450 Fusion Systems

Chapter 3

The Haem Domain of CYP102A3

3. The Haem Domain of CYP102A3

3.1. General Information

As stated in the Introduction, the first redox partner fusion P450 enzyme to be identified was P450 BM3 (CYP102A1) from *Bacillus megaterium*. Thanks to its reductase system resembling mammalian microsomal P450 enzymes, BM3 has become a well-studied bacterial model of these systems[71]. While a number of crystal structures exist for the isolated P450 (haem) domain of BM3[87, 106, 137-144], a structure of the full length enzyme has proved elusive due to an inability to crystallise it. Probable reasons for the difficulty in obtaining crystals of the intact enzyme are the dynamic nature of the multi domain enzyme and the fact that it exists as a dimer in solution[72]. Several homologous enzymes belonging to the CYP102 family have been identified since the discovery of BM3. These currently include twelve in the same subfamily (CYP102A) notated as A2-A13 and other enzymes with less amino acid sequence identity in other subfamilies (CYP102B-H)[81]. Information on these enzymes is presented in Table 3.1. Based on their sizes, it appears that all the majority of CYP102 enzymes are redox partner fusions like BM3 (CYP102A1). The exceptions are the enzymes of the CYP102B and C subfamilies and CYP102H1, which all appear to be conventional P450 enzymes resembling the P450 BM3 haem domain.

CYP102 Enzyme	Species of origin	Size (AA Residues)	First Characterised
A1	<i>Bacillus megaterium</i>	1048	1986 Narhi and Fulco[65]
A2	<i>Bacillus subtilis</i>	1061	2004 Gustafsson et al.[79]
A3	<i>Bacillus subtilis</i>	1054	2004 Gustafsson et al.[79]
A4	<i>Bacillus anthracis str. Ames</i>	1065	
A5	<i>Bacillus cereus</i>	1065	2007 Chowdhary et al.[145]
A6	<i>Bradyrhizobium japonicum</i>	1078	
A7	<i>Bacillus licheniformis</i>	1074	2008 Dietrich et al.[146]
A8	<i>Bacillus thuringiensis</i>	1065	
A9	<i>Bacillus weihenstephanensis</i>	1065	
A10	<i>Erythrobacter litoralis</i>	1070	
A11	<i>Erythrobacter sp. NAP1</i>	1071	
A12	<i>Rhodopseudomonas palustris</i>	1072	
A13	uncultured soil bacterium	1079	
B1	<i>Streptomyces coelicolor</i>	527	2010 Lamb et al.[147]
B2	<i>Streptomyces avermitilis</i>	504	
B3	<i>Rhodococcus sp</i>	498	
B4	<i>Streptomyces scabies</i>	513	
C1	<i>Rhodococcus sp</i>	Partial sequence	
C2	<i>Rhodococcus erythropolis</i>	510	
D1	<i>Streptomyces avermertilis</i>	474	
E1	<i>Cupriavidus metallidurans</i>	1064	
F1	<i>Actinosynnema pretiosum subsp. Auranticum</i>	1005	
G1	<i>Streptomyces scabies</i>	1070	
G2	<i>Saccharopolyspora erythraea</i>	1061	
H1	<i>Nocardia farcinica</i>	473	

Table 3.1. List of known enzymes of the CYP102 family, their length and (where applicable) year of first characterisation and reference. This list is based on the list of P450 enzyme sequences by Nelson[81] along with searches for some partial sequences using NCBI blast, which provided the full sequences.

The first of the CYP102A enzymes to be isolated (following BM3) were CYP102A2 and A3 from *Bacillus subtilis*, both being full fusion systems[79]. They were found to express well in heterologous (*E. coli*) expression systems, but in the case of CYP102A2 without full population of bound haem in the isolated enzyme. CYP102A3 is therefore a stable, BM3 homologue known to be easily expressed and purified and which may be more amenable to crystallisation than is BM3 itself. Additionally, while having activity against a similar range of fatty acid substrates, there are some differences in regioselectivity between the two enzymes. While BM3 favours hydroxylation of myristic acid at the ω -1 position with minor ω -2 and ω -3 hydroxylated products, CYP102A3 favours the ω -2 product[79]. This altered reactivity could be due to structural differences between the active sites of the two enzymes. Indeed, there are differences in the amino acid sequences which indicate the absence in CYP102A3 of residues found to control substrate binding in BM3. Arg-47 and Tyr-51 in BM3 are situated at the mouth of the substrate binding cavity and hydrogen bond with the carboxyl terminus of fatty acid substrates via their side chains [69]. Amino acid sequence alignments show that these residues are unique to BM3, being non-conserved in all other CYP102A enzymes (Figure 3.1). This provides a possible explanation for the tendency of CYP102A3 to hydroxylate further up the fatty acid chain.

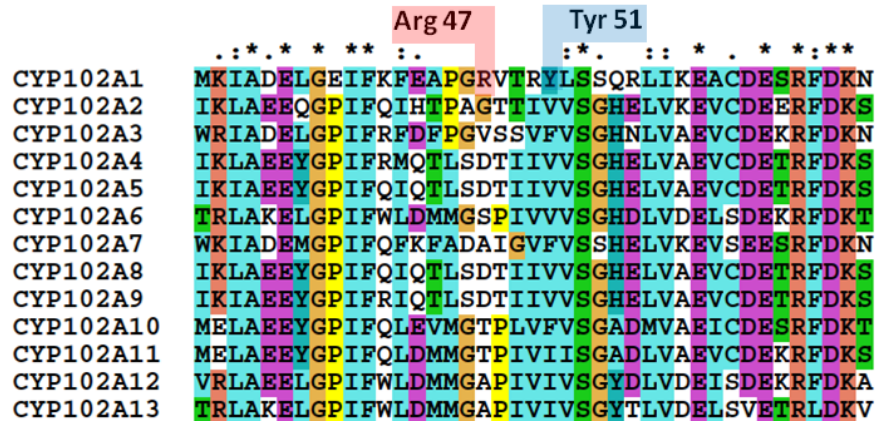


Figure 3.1. Section of an amino acid sequence alignment of members of CYP102A P450 subfamily (residues 30 – 70 relative to A1). The alignment shows the absence of Arg 47 and Tyr 51 residues from all members except BM3 (CYP102A1).

That is to say, if the substrate is not anchored in the same position as is observed in the crystal structures of fatty acid-bound forms of the BM3 haem domain[140, 148], it

may enter further into the substrate binding cavity allowing better access to the haem reaction centre for the ω -2 and ω -3 positions of myristic acid in CYP102A3 and A2. Two further differences between BM3 and CYP102A3 amino acid sequences occur at Ala-74 and Leu-188 in BM3 where CYP102A3 has a glycine and a serine respectively (sequence alignments shown in Figure 3.2).

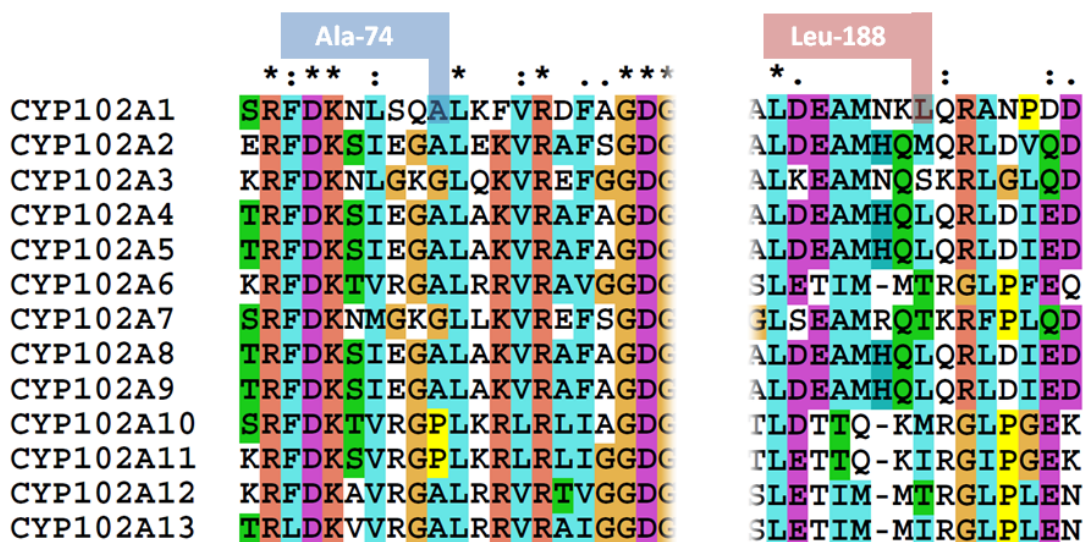


Figure 3.2. Sections of an amino acid sequence alignment of members of CYP102A P450 subfamily. The sections show residues 65 – 85 (left) and 180 – 195 (right) relative to A1. Highlighted in the alignments are the residues Ala-74 and Leu-188 in BM3 (CYP102A1), which are not fully conserved throughout the sub family and in CYP102A3 are altered to glycine and serine respectively.

These residues in BM3 have been mutagenised previously (as part of a A74G/F87V/L188Q triple mutant), in work aimed at increasing activity against aromatic hydrocarbons. This BM3 mutant showed an increase, of 2-4 orders of magnitude, in activity against the aromatic compounds screened (compared to wild type BM3) with positive contributions from the A74G and L188Q mutations[149]. In this work it is suggested that these residues make up the opening of the substrate binding cavity and that this is the origin of their effect on substrate selectivity.

An obvious starting point, therefore, is the structural characterisation of the haem domain of CYP102A3. A haem domain structure would simplify solving of the full length structure and an additional structure with bound substrate could help to determine the effect of differences between the amino acid sequences of CYP102A3 and BM3 upon substrate binding and catalysis. In this way the active-site environment of each enzyme could be compared and a structural basis for the differences in reactivity discovered. The structures of the BM3 haem domain and the

mutagenesis studies that followed provided an insight into the active site properties of not only CYP102s, but all P450 enzymes. Moreover, they showed potential biotechnological applications by allowing the creation of mutant enzymes with altered substrate specificity and/or reactivity. Details of work aimed at engineering aspects of the BM3 reaction by rational mutagenesis and directed evolution are contained in the introduction (section 1.1.10).

The major goal of this section of the project was therefore to determine the crystal structure of the haem domain of CYP102A3. Alongside structural determination, other experiments were carried out to explore binding of ligands and the impact of ligand binding upon stability of the protein and on the redox potential of the haem iron. The theory of the experimental methods used and their relevance to the study of P450s is discussed in the introduction (section 1.2).

Substantial characterisation has already been carried out for the full length CYP102A3 enzyme. Most importantly, substrate binding and turnover kinetics were carried out on a range of fatty acid substrates[79]. These experiments showed overall weaker binding affinities and lower catalytic rates for unbranched fatty acid substrates with CYP102A3 compared to BM3. A selection of branched chain substrates showed not only improved catalytic rates but also apparent cooperativity in binding, with sigmoidal binding and kinetic dependencies in some cases, which were not previously observed for BM3. Additionally, product determination by HPLC and GC/MS revealed differences in regioselectivity between CYP102A3 and BM3 outlined above[79]. Some of these experiments were repeated with HDCYP102A3 to allow comparisons to be made with the full length enzyme. It was hoped that these experiments would highlight the influence of the reductase domain and (likely) the dimeric nature of the full length protein on the biophysical and other properties of the haem domain. These experiments include substrate binding studies to analyse K_d values and binding patterns (*e.g.* cooperativity) and the measurement of redox potentials for the system to compare with BM3 and to examine the influence of substrate binding on the thermodynamic properties of the enzymes.

3.2. Results

3.2.1. Expression and Purification of CYP102A3 haem domain

The haem domain of CYP102A3 (HDCYP102A3) was expressed and purified as described in section 2.4.7. Purity was initially assessed spectrophotometrically by the ratio of haem Soret band ($\lambda \approx 420$ nm) to protein aromatic residues ($\lambda \approx 280$ nm), typically referred to as the Reinheitszahl (Rz) value, a common measure of haemoprotein purity.

Purity of protein species was also assessed by SDS-PAGE. The purified HDCYP102A3 stock was run on a 12% SDS-PAGE gel (Figure 3.3) to confirm purity. Even when overloaded, only one band was observed at a molecular mass of ~ 55 kDa which corresponds to the mass of HDCYP102A3 predicted from its amino acid sequence (54.2 kDa).

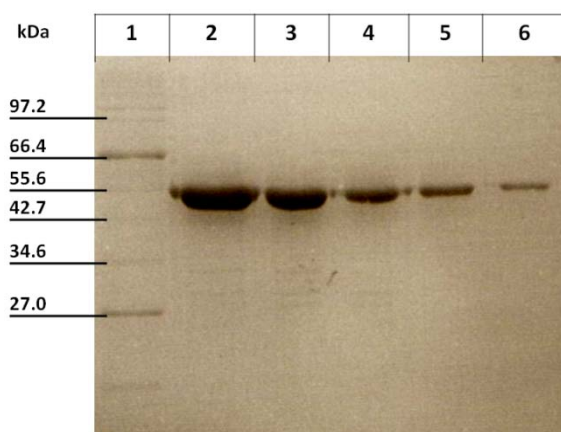


Figure 3.3. Image of an SDS-PAGE gel (12%) showing samples of the haem domain of CYP102A3 at varying concentrations. Lanes are as follows: 1. Broad range protein marker (New England Bioscience) weights are labelled on the left. 2-6. HDCYP102A3 16, 8, 4, 2 and 1 μ g, respectively.

The presence of a P450 protein was confirmed by formation of the reduced carbonmonoxyhaem complex as described in 2.4.11. The resulting spectra (Figure 3.4) show a gradual decrease in absorbance at 418 nm and a corresponding increase at 450 nm following introduction of CO and reductant to the P450 sample. This indicates formation of a P450 complex where the haem thiolate ligation is maintained

while CO binds to the ferrous haem as a sixth ligand. In some P450s (*e.g.* CYP121A1 and CYP51B1 from *M. tuberculosis*) the cysteine thiolate ligand can become fully or partly protonated to a thiol upon reduction and/or reduction/CO binding, giving rise to a Soret absorbance at ~420 rather than ~450 nm. This is not seen in the spectra for CYP102A3, indicating thiolate ligand retention upon reduction[150]. The α/β features also appear to merge into a single peak at 549 nm, an indicator of haem reduction. The key to the slow formation of the P450 complex (several minutes for completion) is that haem reduction by sodium dithionite is relatively slow and only a small proportion of the haem is initially reduced (indicated by the minimal spectral change upon dithionite addition). This small population of reduced haem can readily form the carbonmonoxy complex, and over time the oxidised/reduced equilibrium is pulled in favour of the reduced haem species as a consequence of it being trapped in the stable CO-bound form. The alternative scenario is that the ferrous haem binds oxygen, which then leads to an abortive release of superoxide (in most cases) and the restoration of the ferric haem iron.

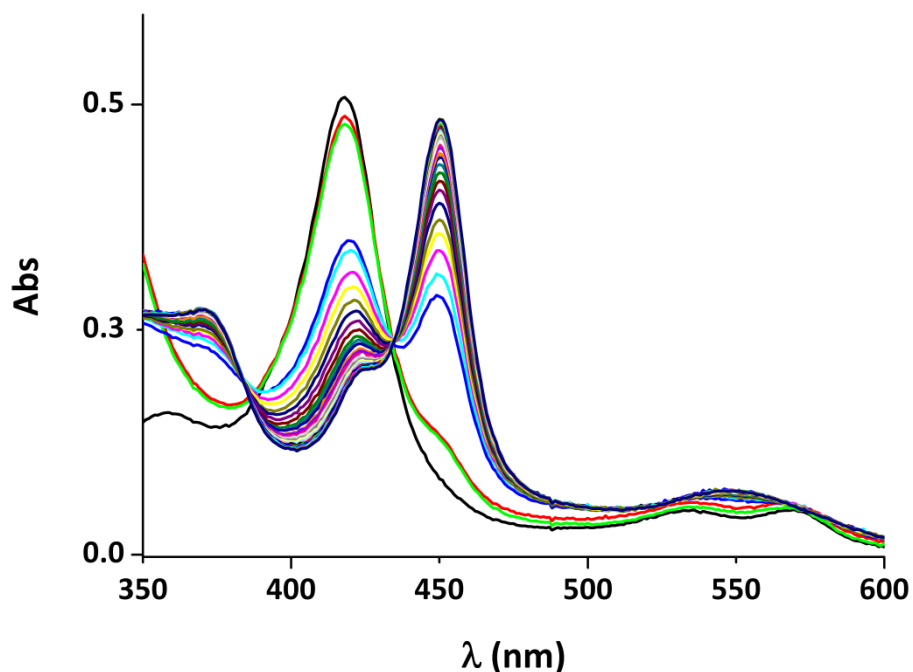


Figure 3.4. UV/Visible spectra showing the changes in absorbance upon reduction and binding of CO to the iron of the haem domain of CYP102A3. The spectrum in black is HDCYP102A3 at the start of the experiment (5 μ M). Spectra in red and green are after the addition of the reducing agent sodium dithionite (1 mg). All other spectra represent a time course with spectra recorded at one minute intervals after carefully mixing the enzyme solution with gaseous CO in a syringe. The spectra show gradual formation of the carbonmonoxy haem (P450) complex and the corresponding characteristic signal at 450 nm. An isosbestic point is clearly visible at 434 nm, representing the overlapping spectra for the ferric haem and the ferrous-CO species.

3.2.2. Ligand Binding Studies

Ligand binding titrations were carried out in order to probe substrate preference and inhibitor binding in HDCYP102A3. Full length CYP102A3 has been shown to differ from its close homologue BM3 with respect to substrate specificity and the regioselectivity of substrate hydroxylation, as discussed previously (section 3.1).

The ligands for P450 enzymes can be broadly broken down into substrates and inhibitors based on the types of optical shifts they induce. Most P450 substrates bind in the active site near the haem, displacing the water ligand from the haem iron and promoting a high-spin configuration of the ferric haem iron *d*-electrons. The majority of P450 inhibitors are those which bind the haem iron as a sixth ligand and promote or reinforce a low-spin state. With each of these binding modes there is a perceivable

change in the haem Soret band in the UV/Visible spectrum, usually seen as a shift to a shorter wavelength for substrates (type I shift) and to a longer wavelength for haem-coordinating inhibitors (type II shift)[43] (see section 1.1.6).

Binding titrations were carried out as described in section 2.4.12, in assay buffer. Examples of the spectra resulting from a substrate and an inhibitor titration are shown (Figure 3.5 & Figure 3.10). The initial spectrum of HDCYP102A3 recorded in the absence of ligand was subtracted from each spectrum recorded after an addition of ligand to create a difference spectrum at each point in the titration (Figure 3.6 & Figure 3.11). From the difference spectra the wavelengths where the maximum change in absorbance occurred were selected, and the difference in absorbance between these wavelengths computed ($\Delta\Delta\text{Abs}$). These data were then plotted against the concentration of ligand and, according to the shape of the data, could be fitted to a binding function to derive a value of the dissociation constant K_d (Figure 3.7, Figure 3.8 & Figure 3.12). The binding functions used were outlined previously (section 1.2.2, Equation 1.4Equation 1.5).

3.2.2.1. Substrates

As CYP102A3 is a fatty acid hydroxylase enzyme [79], a range of saturated fatty acids of varying chain lengths (C_{12} - C_{18}) were chosen to probe binding preferences of HDCYP102A3. These are lauric acid (C_{12}), myristic acid (C_{14}), palmitic acid (C_{16}) and stearic acid (C_{18}). These substrates were selected based on those used in previous work on BM3 and full length CYP102A3[79, 151]. A selection of branched chain fatty acids were used, as previous work had revealed possible cooperative binding and kinetic dependencies with CYP102A3 holoenzyme[79]. These were 12- and 13-methylmyristic acids, 14-methylpentadecanoic acid and 15-methylpalmitic acid. Lastly, two further substrates shown to have a high binding affinity and catalytic rate with BM3 were used. These were arachidonic acid – a long chain (C_{20}) polyunsaturated fatty acid[69] and *N*-palmitoyl glycine (NPG) – an amino acid derivatised fatty acid and tight binding substrate for BM3[148].

A sample set of spectra for 13-methylmyristic acid are presented in Figure 3.5, and are typical of most sets of spectra recorded for substrates. The corresponding difference spectra are shown in Figure 3.6, and are similarly representative of difference spectra for other substrates.

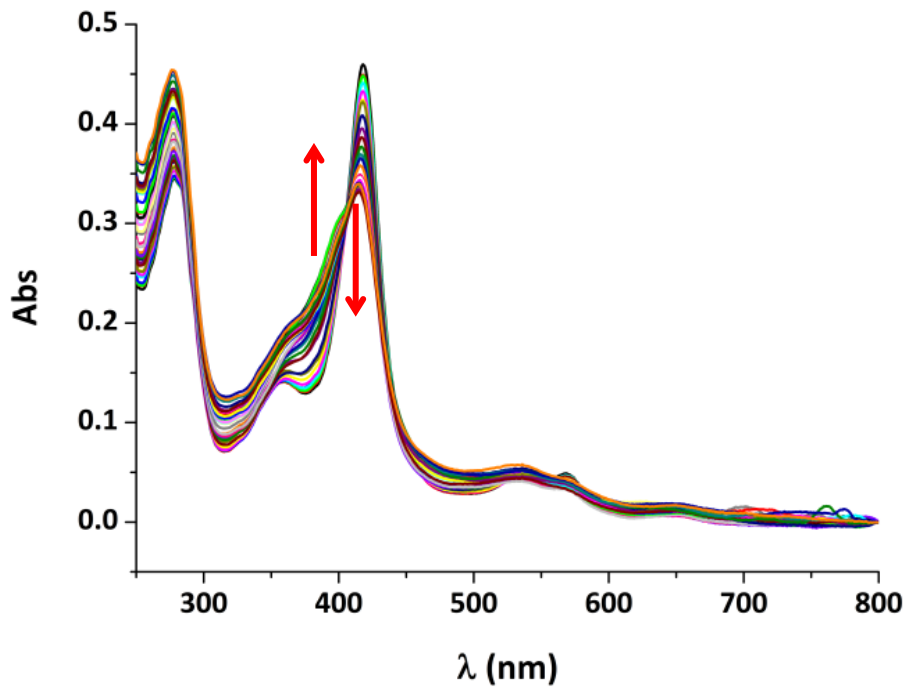


Figure 3.5. UV/visible spectra of the titration of 13-methylmyristic acid (5 - 130 μM) against HDCYP102A3 (5 μM). The UV/Visible spectrum (between 250 and 800 nm) wavelength was recorded initially and then after each addition of ligand. All spectra are baselined at 800 nm. Red arrows indicate direction of absorbance changes in the Soret band region on substrate addition.

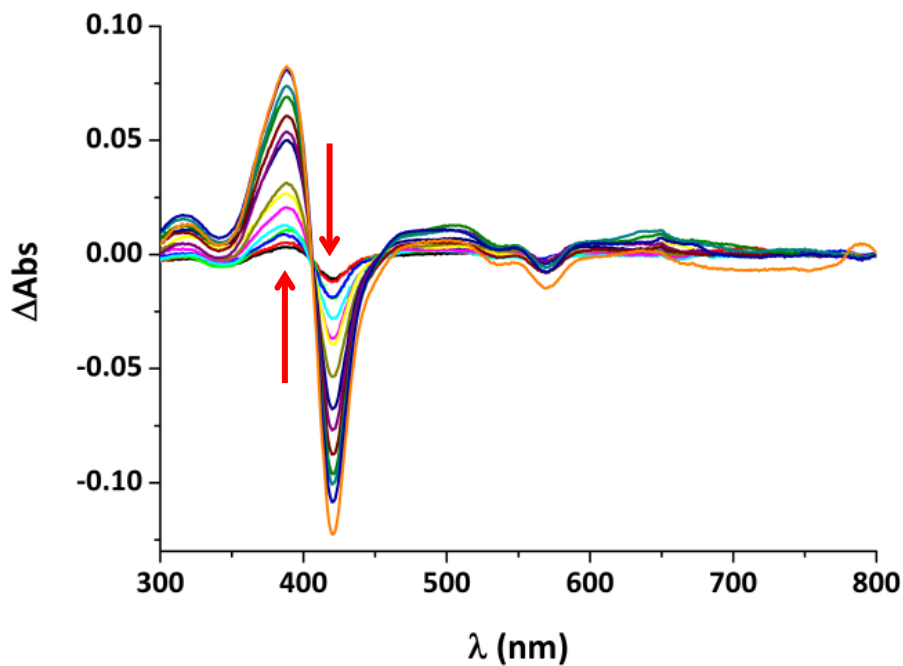


Figure 3.6. Difference spectra for the titration of 13-methylmyristic acid (5 - 130 μM) against HDCYP102A3 (5 μM). The initial spectrum of HDCYP102A3 in the absence of ligand was subtracted from each of the spectra recorded after sequential addition of the substrate to create difference spectra. This highlights the shift in haem Soret absorbance band (Red arrows indicate the direction of the shift). These spectra show a type I shift with a trough and a peak forming at 421 nm and 388 nm respectively. An isosbestic point is visible at 406 nm.

Substrate binding to HDCYP102A3 was seen to cause a type I shift of the haem Soret band to a shorter wavelength (384 – 389 nm) with a small variation in the wavelength depending on the specific substrate. This shift is caused by displacement of the water molecule, which is the sixth ligand to the haem iron, upon substrate binding, leading to a pentacoordinate ferric haem. This also promotes a high-spin electronic configuration (see section 1.1.6). Binding curves for substrates are presented in Figure 3.7 (one-site hyperbolic fit) and Figure 3.8 (sigmoidal (Hill) fit).

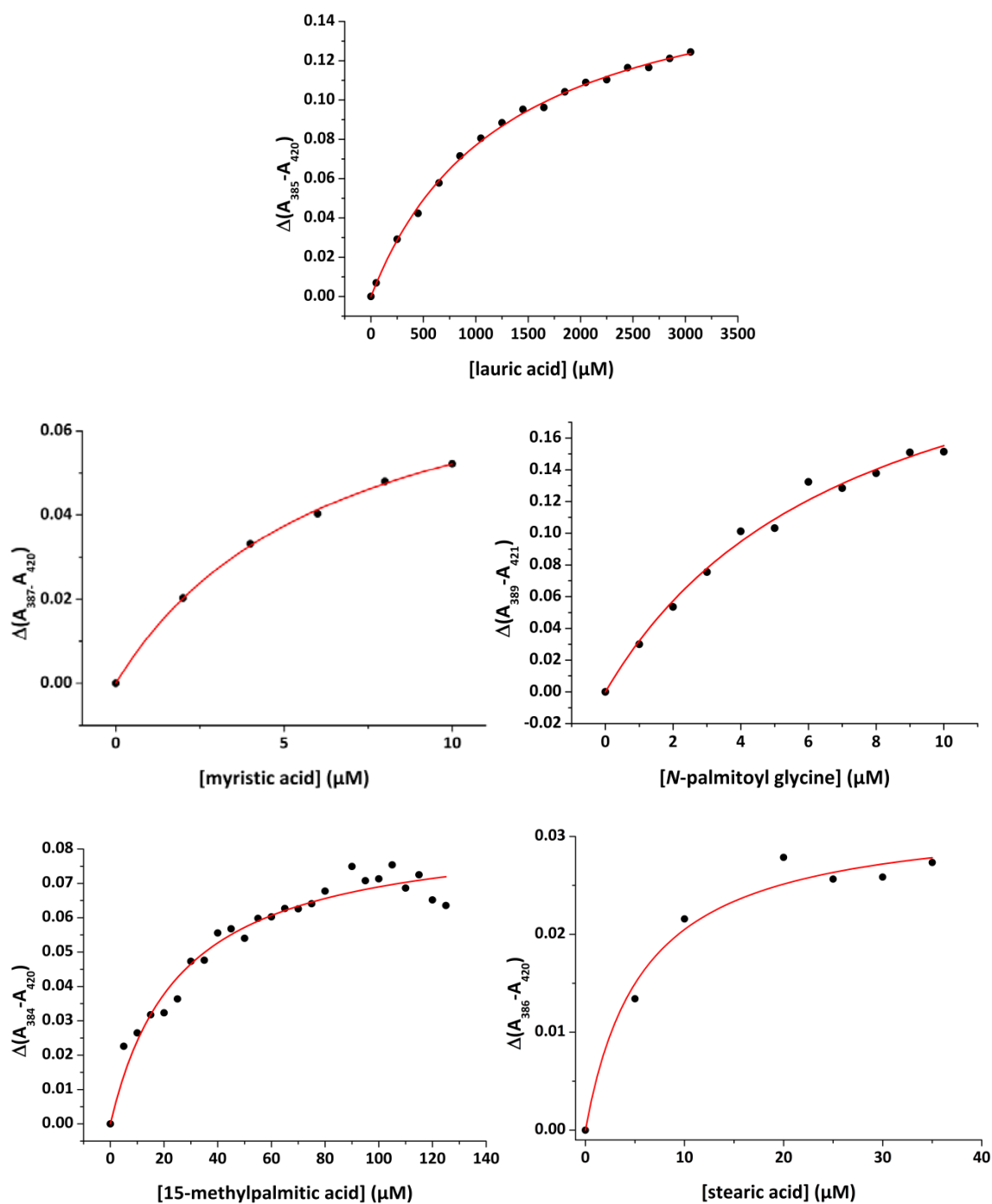


Figure 3.7. Binding curves for HDCYP102A3 ($\sim 5 \mu\text{M}$) against a selection of fatty acid substrates. Changes in absorbance associated with a shift in the haem Soret signal are plotted against the concentration of substrate for each addition in the titration. The data are fitted using a hyperbolic one-site binding function described earlier (section 1.2.2, Equation 1.4) to produce values of the dissociation constant (K_d). Fits are shown in red.

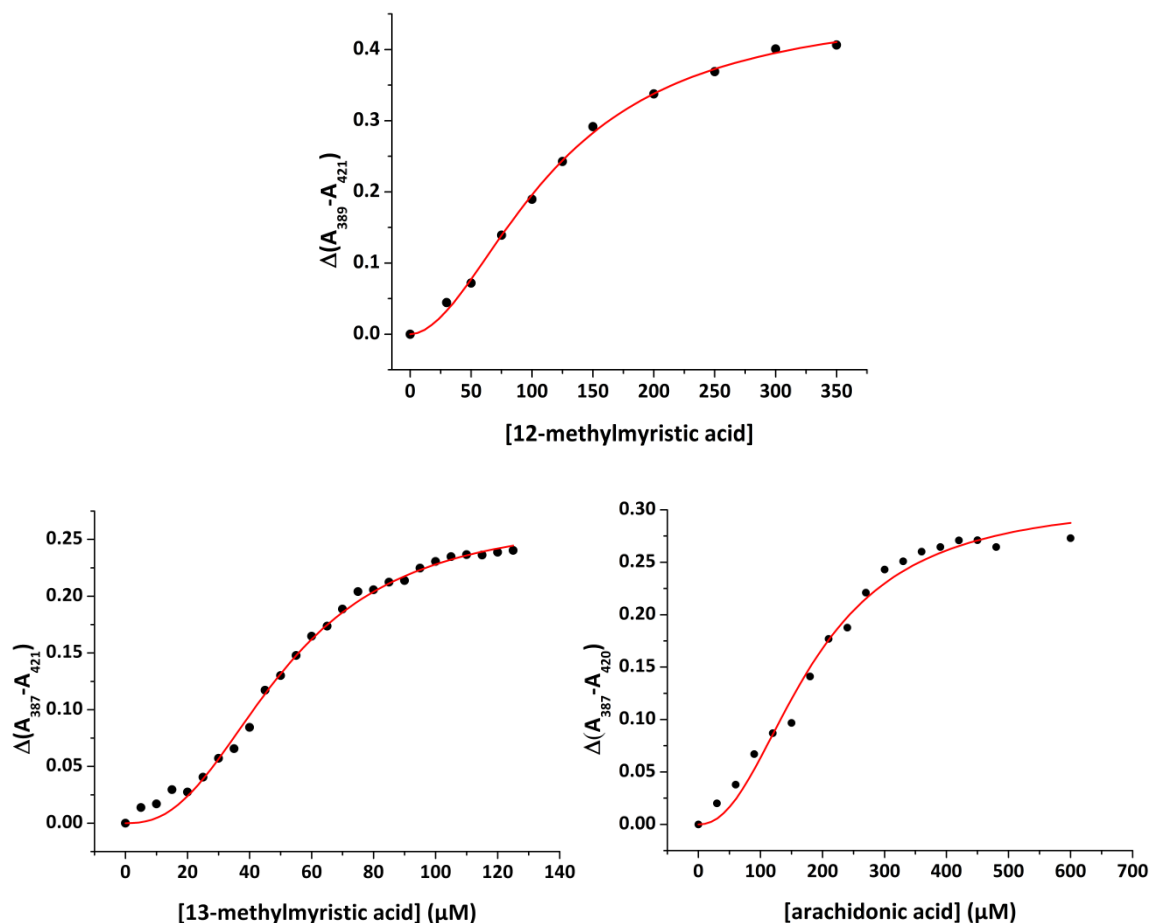


Figure 3.8. Binding curves for HDCYP102A3 ($\sim 5 \mu\text{M}$) against three fatty acid substrates: 12-methylmyristic acid (20-350 μM), 13-methylmyristic acid (5-125 μM) and arachidonic acid (30-600 μM). Changes in absorbance associated with a shift in the haem Soret signal are plotted against the concentration of substrate for each addition in the titration. The data were then fitted using the Hill function described earlier (section 1.2.2, Equation 1.5) to produce values of the apparent dissociation constant (K_{H}). These values are presented in Table 3.2.

The results show an observable binding interaction for HDCYP102A3 in the presence of a range of fatty acid substrates (see Table 3.2). In several cases the binding curve appears incomplete and this is due to substrate insolubility at higher concentrations. Upon adding substrate from a 50% ethanol solution to the aqueous buffer its solubility was somewhat reduced. At higher concentrations this leads to turbidity which interfered with collection of the absorbance spectra. In these cases the plateau of the binding curve (A_{max}) is inferred in the fit and may not reflect the true limit of binding.

Sigmoidal binding curves were observed for 13-methyl myristic and arachidonic acids, as seen for the full length enzyme [79]. When fitted to a Hill function they both showed a Hill number (n) of ~ 2.2 . When this binding behaviour was seen in the full length enzyme by Gustafsson *et al.* it was commented that the cooperativity could be due to the dimeric nature of the enzyme in solution [79]. However, the isolated haem

domain of CYP102A3 has been shown to be monomeric in solution by the light scattering experiments (see below, 3.2.8). Two possible explanations for the same effect are discussed in recent work on CYP130A1 from *M. tuberculosis*[152]. Firstly, the cooperative binding may arise from multiple molecules of substrate binding to the same enzyme active site. Cooperative binding of this nature has been demonstrated before in P450s such as P450EryF (CYP107A1)[153], CYP3A4 in humans[154] and CYP158A2[155]. In this case resultant Hill numbers were ~ 1.3 compared to $\sim 1.8 - 2.5$ for the HDCYP102A3 examples.

Another possibility for cooperative binding is the substrate-induced dimerisation of HDCYP102A3. This effect has been crystallographically shown to occur for a P450 unit in the work mentioned above on CYP130A1 from *M. tuberculosis*. The researchers postulated that the binding of substrate to a monomer would cause a conformational change in the enzyme that promotes dimerisation[152]. A substrate-bound structure might also confirm that the same phenomenon occurs in HDCYP102A3.

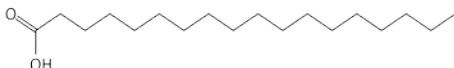
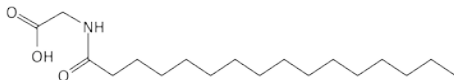
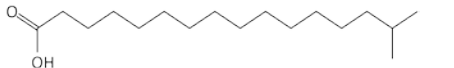
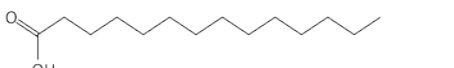
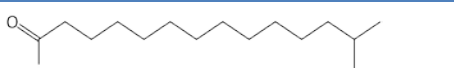
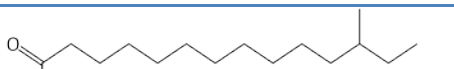

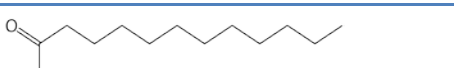
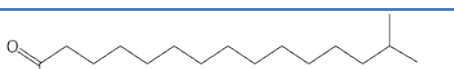
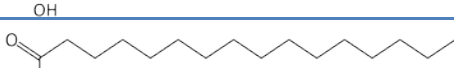
Substrate		K_d/K_H^* (μM)	n
Stearic acid		6.0 \pm 2	
<i>N</i> -palmitoyl glycine		7.4 \pm 1.2	
15-methylpalmitic acid		26 \pm 3.5	
Myristic acid		26 \pm 3.4	
13-methylmyristic acid		46 \pm 2*	2.5 \pm 0.2
12-methylmyristic acid		118 \pm 5*	1.8 \pm 0.1
Arachidonic acid		185 \pm 12*	2.0 \pm 0.2
Lauric acid		1670 \pm 263	
14-methylpentadecanoic acid		n.a.	
Palmitic acid		n.a.	

Table 3.2. Dissociation constants for HDCYP102A3 with a range of fatty acid substrates. K_d values were determined by fitting the experimentally acquired binding curves using a one site binding equation, except in the case of arachidonic and 12/13-methylmyristic acids. These substrates produced binding curves with sigmoidal character and these were fitted using the Hill function. In these cases an observed dissociation constant ' K_H ' is shown (indicated *) and the Hill number (n) is given.

Table 3.2 shows a general pattern that saturated substrates have a higher affinity according to their chain length. The substrate with the lowest K_d was stearic acid (C_{18}) while the substrate with the highest K_d was lauric acid (C_{12}). The branched substrate 14-methylpentadecanoic acid (C_{15}) did not appear to bind at all. Arachidonic acid is polyunsaturated and as such may interact differently with the substrate binding cavity of HDCYP102A3. This pattern of affinity is similar to that observed for BM3 in that lauric acid shows the highest K_d value (100 μM)[151], although this is an order of magnitude lower than the value for HDCYP102A3. The saturated substrate with the lowest K_d value for BM3 is myristic acid (C_{14}). In the case of arachidonic acid, the substrate producing the highest catalytic rate with BM3, the K_d value recorded for HDCYP102A3 (18.5 μM) is around 5 times that previously

measured for BM3. This might be what is reflected, although not proportionately, in the tenfold difference in the k_{cat} value for this substrate between BM3 (17100 min⁻¹)[69] and CYP102A3 (1690 min⁻¹)[79].

3.2.2.2. Inhibitors

For inhibitors, the common haem ligand imidazole was used along with the derivatives 1-, 2- and 4-phenylimidazoles, which are known P450 inhibitors with higher binding affinities than imidazole (Figure 3.9). Imidazole is an electronically polar molecule which hinders its access to the non-polar substrate cavity of a P450 and can make it a relatively weak binding ligand. The phenyl group of the phenylimidazoles decreases the overall polarity of the molecules allowing them better access to the substrate cavity and the haem. Additionally, with the phenyl group at various positions on the imidazole ring, the availability of N3, the ligating nitrogen, is varied. This is reflected in the relative binding strength of these ligands.

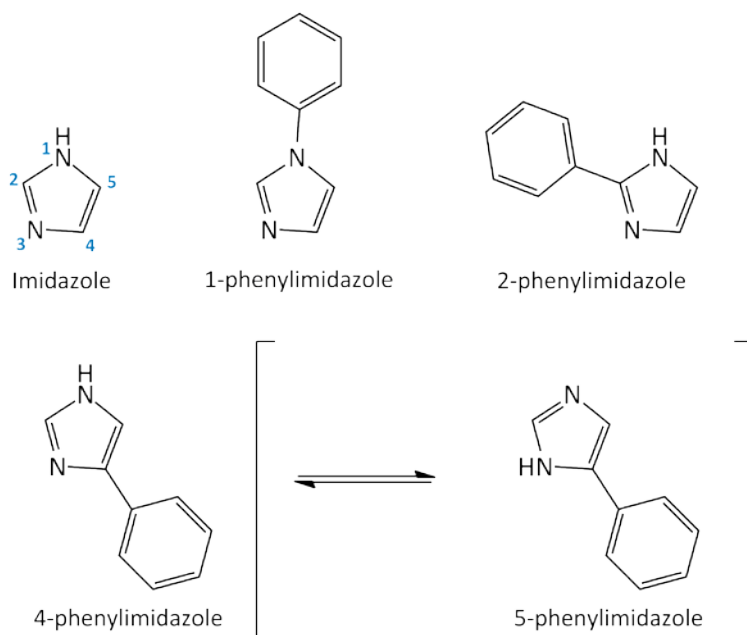


Figure 3.9. Structures of azole inhibitors used for binding titration experiments. The numbering system for imidazole atoms is shown on the structure of imidazole. Tautomerism between 4- and 5-phenylimidazoles is shown as this may be relevant to the binding mode of this inhibitor.

In the case of 1-phenylimidazole the phenyl group is attached to N1, two bonds away from N3, and is not conjugated with the imidazole ring. The overall polarity of the molecule is reduced compared to other phenylimidazoles as the N-C₆H₅ bond is less polar than an N-H bond.

2-phenylimidazole has the phenyl group positioned on C2, next to N3 and conjugated to the aromatic system of the imidazole ring. The proximity of the phenyl group to the ligating nitrogen causes steric hindrance and reduces binding affinity for this inhibitor.

4-phenylimidazole most likely exists in solution as a mixture of tautomers with a majority of 4- and a small proportion of 5-phenylimidazole (3-9%)[156] (Figure 3.9). In the latter the phenyl group would be two bonds from N3, as in 1-phenylimidazole, allowing better access for the haem iron, although the phenyl group would still be conjugated with the imidazole ring. If 5-phenylimidazole is the ligand that binds, the equilibrium could be pulled in favour of this tautomer. Indeed, in the P450cam structure with 4-phenylimidazole bound, the 5-phenyl form is shown[157].

A sample set of spectra for 4-phenylimidazole binding are presented in Figure 3.10, and are typical of most sets of spectra recorded for inhibitors. The corresponding difference spectra are shown in Figure 3.11, and are similarly representative of difference spectra for other inhibitors.

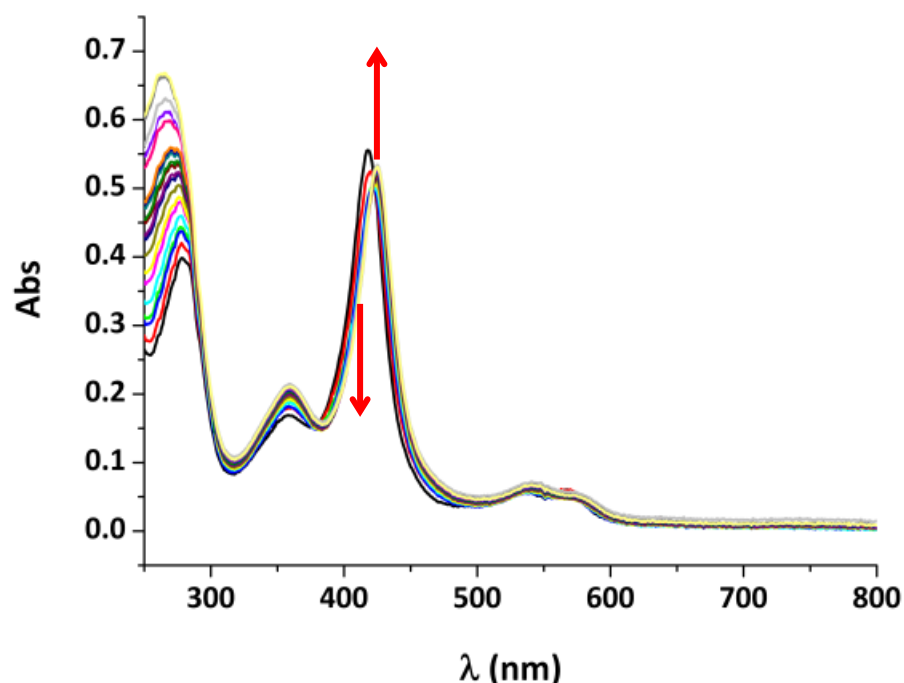


Figure 3.10. UV/visible spectra of the titration of 4-phenylimidazole (5 - 100 μ M) against HDCYP102A3 (5 μ M). The UV/Visible spectrum (between 250 and 800 nm) wavelength was recorded initially, and then after each addition of ligand. All spectra are baselined at 800 nm. Red arrows indicate direction of absorbance changes in the Soret band region on inhibitor addition.

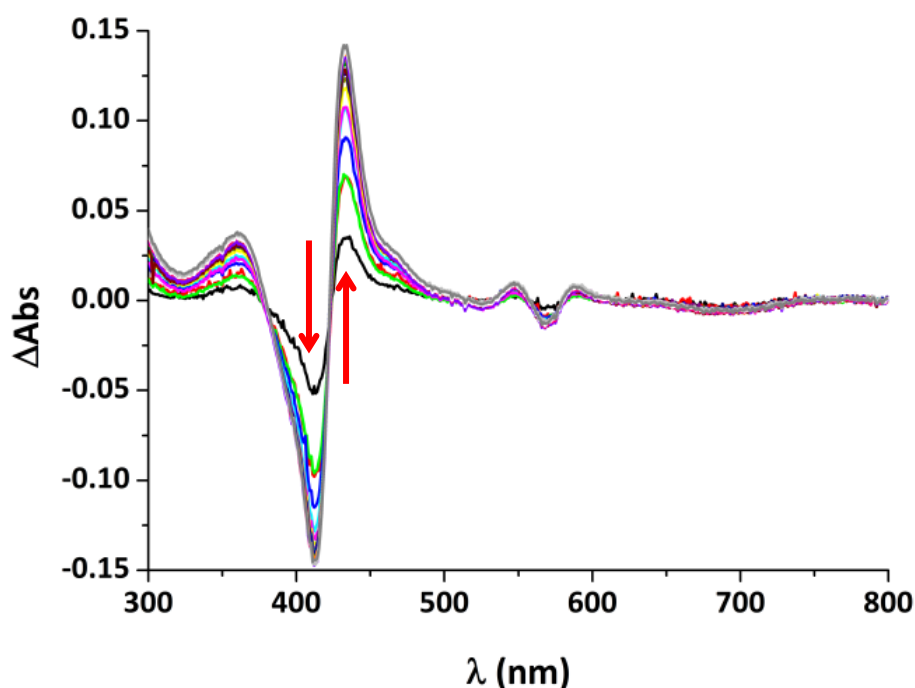


Figure 3.11. Difference spectra for the titration of 4-phenylimidazole (5 - 100 μM) against HDCYP102A3 (5 μM). The initial spectrum of HDCYP102A3 in the absence of ligand was subtracted from each of the spectra recorded after sequential addition of the ligand to create difference spectra. This highlights the shift in haem Soret absorbance band (Red arrows indicate the direction of the shift). These difference spectra show a type II shift with a trough and a peak forming at 413 nm and 434 nm respectively. An isosbestic point is visible at 422 nm.

For azole inhibitors a type II shift to a longer wavelength (431 - 435 nm) was observed, again with a variation in the wavelength depending on the inhibitor. This shift indicates that a spectrochemically stronger ligand has replaced water as a ligand to the haem iron (see section 1.1.6). The plotted binding curves (Figure 3.12) show a hyperbolic dependence on inhibitor concentration (as is common for binding of imidazole to P450 haem) in each case, except that of 1-phenylimidazole. The binding data for this ligand are fitted poorly using a hyperbola, but can be fitted well using a quadratic binding function (section 1.2.2, Equation 1.6) that is usually appropriate in cases of tight binding ligands.

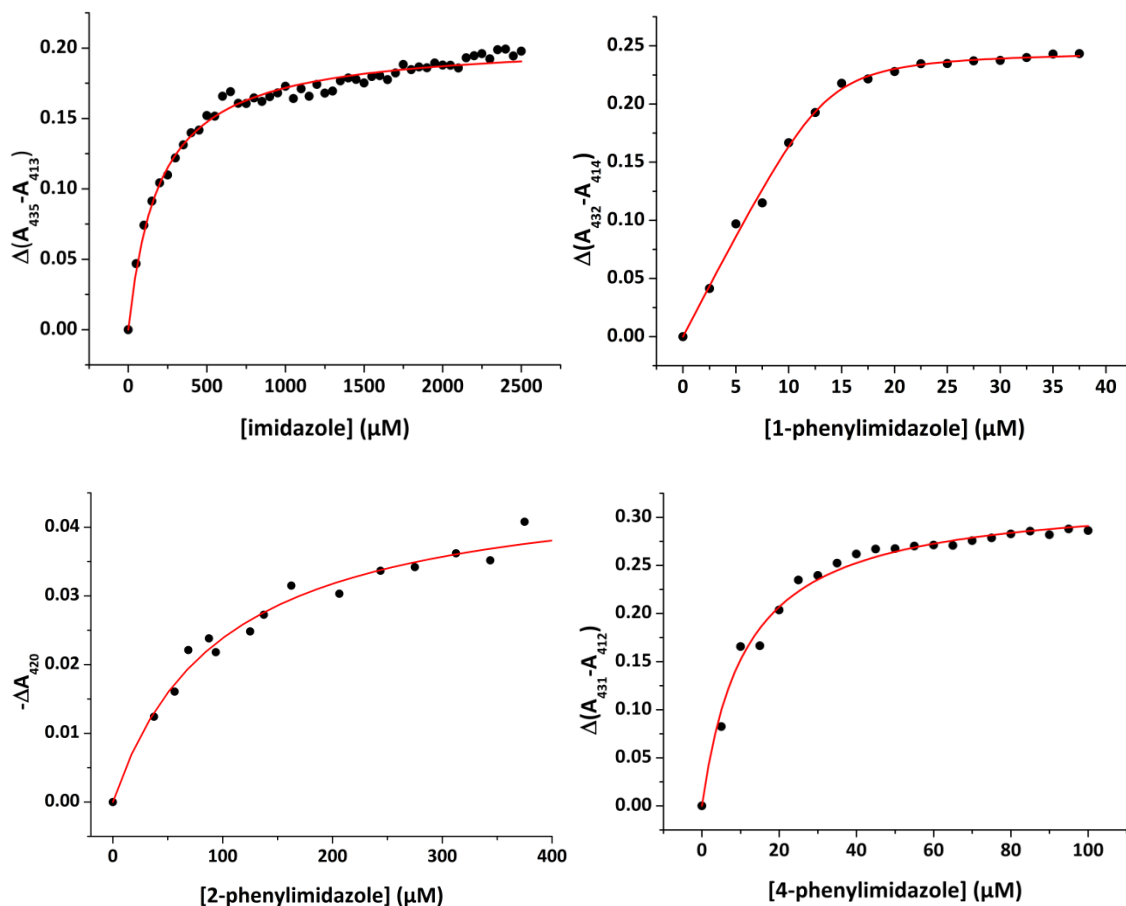


Figure 3.12. Binding curves for a selection of inhibitors titrated against CYP102A3. Imidazole and 1-phenylimidazole (top), 2-phenylimidazole and 4-phenylimidazole (bottom). Data were derived by calculating the difference between the Δ Absorbance values at the peak and the trough in difference spectra, except in the plot for 2-phenylimidazole where the difference spectra showed no obvious developing peak. In this case negative values of Δ Absorbance in the trough at 420 nm were used. All data are fitted using a hyperbolic one site binding function (section 1.2.2, Equation 1.4), except for the data for 1-phenylimidazole, which are fitted using a quadratic (tight binding) function (section 1.2.2, Equation 1.6). The fits provide values of the dissociation constant (K_d). These values are 196 μ M (imidazole), 0.5 μ M (1-phenylimidazole), 107 μ M (2-phenylimidazole) and 11 μ M (4-phenylimidazole). The K_d and $K_{d(app)}$ values are tabulated below with errors and equivalent values for BM3 for comparison (Table 3.3).

The values of the dissociation constants (K_d) for the azole compounds are comparable to those recorded for full length BM3 and CYP102A3. Firstly imidazole itself shows approximately a halving of K_d value for HDCYP102A3 (196 \pm 9 μ M) compared to BM3 (535 \pm 5 μ M)[143] and FLCYP102A3 (513 \pm 30 μ M)[79]. This could suggest that the active site is more accessible to small polar molecules in HDCYP102A3.

Indeed, the value of the apparent dissociation constant for 1-phenylimidazole (0.5 \pm 0.2 μ M) shows a twenty-fold decrease in HDCYP102A3 compared to the value for BM3 (10.4 \pm 0.4 μ M)[69], indicating twenty times greater affinity for this ligand in HDCYP102A3 compared with BM3. A possible explanation for the higher affinity is that aromatic residues in the haem cavity of HDCYP102A3 (*e.g.* Phe-88) are better

positioned to participate in aromatic stacking interactions with the phenyl group of 1-phenylimidazole than are those in BM3. The value of K_d for HDCYP102A3 also differs from that recorded for FLCYP102A3 ($0.95 \pm 0.1 \mu\text{M}$)[79] by $0.15 - 0.75 \mu\text{M}$ (within error limits). This could suggest that the presence of the reductase domain effects a slight reduction in the binding affinity of CYP102A3 for 1-phenylimidazole.

Conversely, in the case of 4-phenylimidazole there is a ten-fold increase in the K_d value for HDCYP102A3 ($11 \pm 1 \mu\text{M}$), compared to that for BM3 ($0.85 \pm 0.45 \mu\text{M}$). This indicates that 4-phenylimidazole is around ten times less tightly bound in HDCYP102A3.

In BM3, 2-phenylimidazole is reported not to show any binding (*i.e.* a lack of a haem optical shift), although such binding phenomena have been shown in other P450s, such as P450cam. In this case a high resolution crystal structure shows 2-phenylimidazole binding indirectly to the haem through a water molecule[157] (RCSB PDB id: 1PHE). The data presented here, for 2-phenylimidazole binding to HDCYP102A3, come only from a reduction in the Soret band absorbance at 420 nm, with no discernable shift in wavelength. This may be explained if this ligand binds in the same way as shown in the P450cam structure, as indirect binding could decrease the spectral change observed. The calculated value of the dissociation constant ($107 \pm 23 \mu\text{M}$) is lower than that for imidazole, which would seem to indicate that 2-phenylimidazole is a better ligand for HDCYP102A3. Despite this, the K_d value is still higher than those recorded for 1- and 4-phenylimidazoles.

All K_d values for inhibitors are presented in Table 3.3 below.


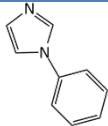
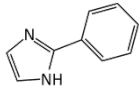
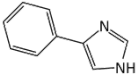
Ligand		K_d (μM)		
		HDCYP102A3	CYP102A3	CYP102A1(BM3)
Imidazole		196 ±9	513 ±30	535 ±5
1-phenylimidazole		0.5 ±0.2	0.95 ±0.1	10.4 ±0.4
2-phenylimidazole		107 ±23*		<i>n.a.</i>
4-phenylimidazole		11 ±1		0.85 ±0.45

Table 3.3. Dissociation constants for HDCYP102A3 with several inhibitor ligands. K_d values were determined by fitting binding curves (shown above) using a one site binding equation (1.2.2, Eqn. 1.4). The binding curve for 1-phenylimidazole was fitted using a tight binding (quadratic) function (section 1.2.2, Equation 1.6). Error values presented are those associated with the fit. Values for the same ligands in BM3 and full length CYP102A3 are presented for comparison, except 2-phenylimidazole for which no binding is reported[69, 143, 151]. The asterisk beside the K_d value for 2-phenylimidazole indicates that this value is based on spectral diminution induced by the ligand, as opposed to a more typical type II spectrum induced on binding the other ligands listed.

3.2.3. Electron Paramagnetic Resonance (EPR) spectroscopy

As stated in the introduction, EPR is used to characterise redox cofactors with unpaired electrons, including ferric P450 haem (section 1.2.4). EPR was performed on HDCYP102A3 to compare the haem environment with those of other P450s.

A sample of HDCYP102A3 for EPR was prepared as described in section 2.4.17.1. The EPR spectrum of HDCYP102A3 was recorded as described in section 2.4.17.2. The EPR experiment was carried out by Dr. Myles Cheesman in the School of Chemical Sciences at the University of East Anglia.

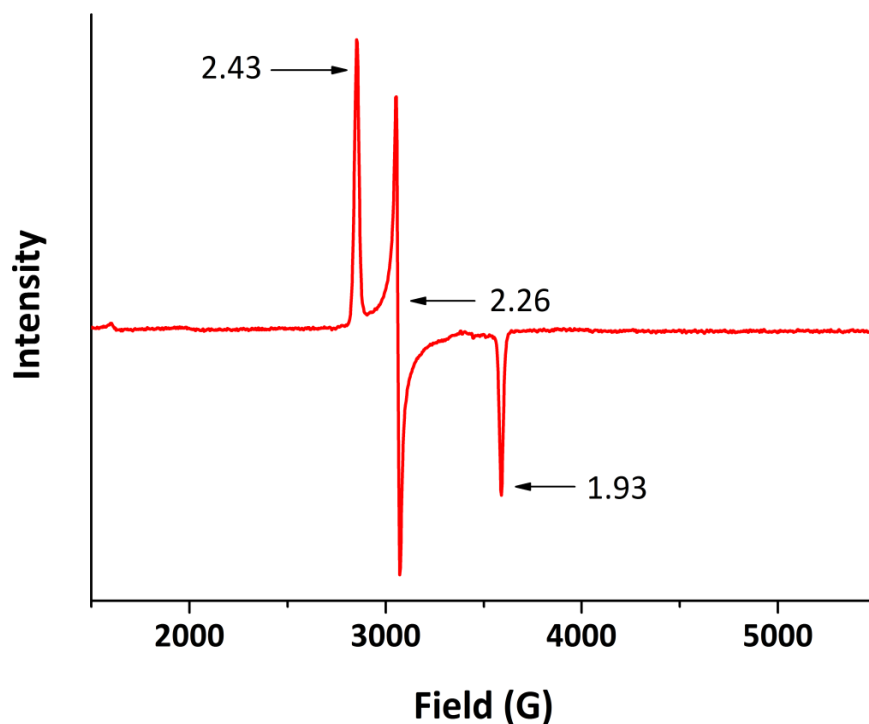


Figure 3.13. EPR spectrum of the HD CYP102A3 (200 μ M) in the oxidised state recorded at 10K. The EPR spectrum was collected by Dr. Myles Cheesman, University of East Anglia. The spectrum shows a rhombic trio of signals characteristic of P450 haem. The associated g-values (labelled on the spectrum) are $g_z = 2.43$, $g_y = 2.26$ and $g_x = 1.93$.

The resulting EPR spectrum (Figure 3.13) shows the classic rhombic trio that has been identified many times in the case of a low spin cysteinate ligated haem. The g-values of these signals are 2.43, 2.26 and 1.93, which are practically identical to those previously seen for wild type BM3 in the low-spin oxidised state (see Table 3.4). The g_z signal at low field with a g-value of 2.43 suggests an O ligand (H_2O or $\cdot\text{OH}$) is coordinated to the sixth position of the haem iron [158].

These g-values, along with those recorded for a selection of other P450 enzymes (including BM3), are presented in Table 3.4 for comparison. The indication of a single low spin population of haem cannot be taken as representative of the physiological enzyme (or indeed the isolated sample) as the EPR spectrum is recorded at a temperature of approximately 10 K. At this temperature the high-spin/low-spin equilibrium of the haem is pushed in favour of the low-spin configuration. Therefore, all of the haem in the sample is quite likely to be low-spin at this temperature regardless of the situation at room temperature[159]. However, absorbance spectra

(e.g. Figure 3.4) do indicate that the HDCYP102A3 is predominantly low-spin at room temperature.

P450 Enzyme	g_z	g_y	g_x
CYP101A1 (P450_{cam})	2.45	2.26	1.91
CYP102A1 (P450_{BM3})	2.42	2.26	1.92
CYP102A3	2.43	2.26	1.93
CYP121A1	2.47	2.25	1.90
CYP51B1	2.42	2.26	1.91

Table 3.4. g -values from electron paramagnetic resonance spectroscopy studies of various P450 enzymes. The characteristic signal resulting from a low-spin haem consists of a rhombic trio of g -tensor elements designated x , y and z in accordance with axes of the molecular orbitals responsible. Data for other P450s are quoted from the literature. CYP121A1 is a C-C bond forming enzyme acting on the cyclic dipeptide cylo-L-Tyr-L-Tyr and CYP51B1 is a sterol demethylase. Both are from *Mycobacterium tuberculosis* [38, 158, 160, 161].

3.2.4. Resonance Raman Spectroscopy

Resonance Raman spectroscopy was carried out on HDCYP102A3 to further characterise the haem environment in terms of haem spin- and oxidation state, conformation of the macrocycle and its substituent groups, and the effect of substrate binding upon these features. As discussed in the introduction, this technique can be applied effectively to haemoproteins by exciting around the wavelength of the Soret absorbance (section 1.2.3).

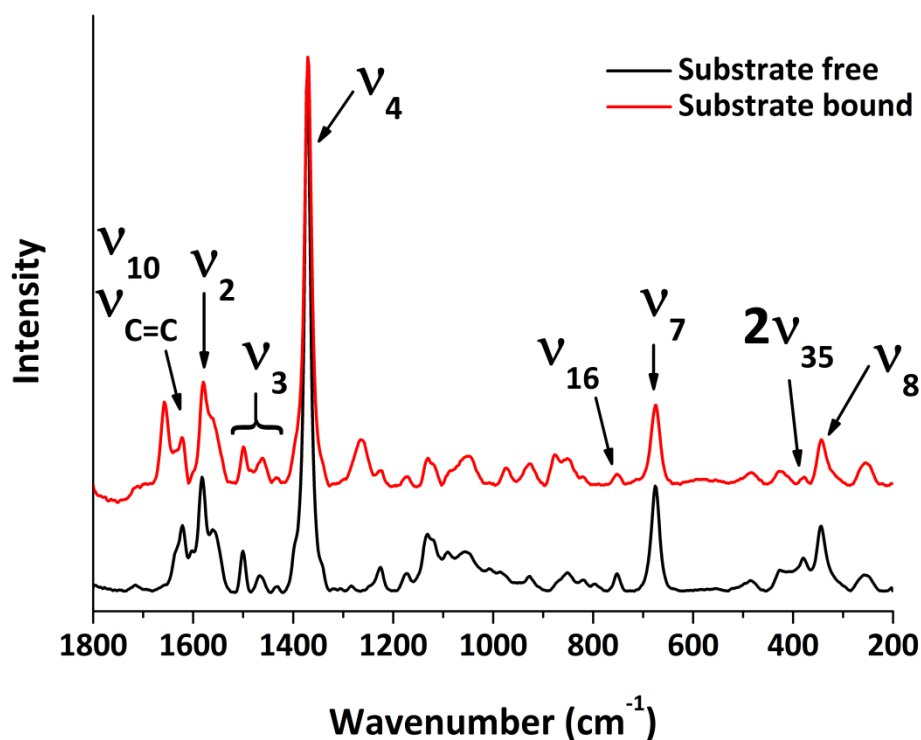


Figure 3.14. Resonance Raman spectra of HDCYP102A3. The spectrum in red is of HDCYP102A3 ($\sim 200 \mu\text{M}$) in the presence of substrate arachidonic acid ($\sim 1 \text{ mM}$) and in black is the substrate free spectrum. Labelled are the ν_4 band and the ν_3 band, which is split into two bands corresponding to low-spin and high-spin electronic configurations of the ferric haem iron. When the substrate is bound, the proportion of low-spin haem decreases, as evidenced by the change in the relative intensities of the ν_3 features. The spectra are displaced on the y-axis for ease of viewing.

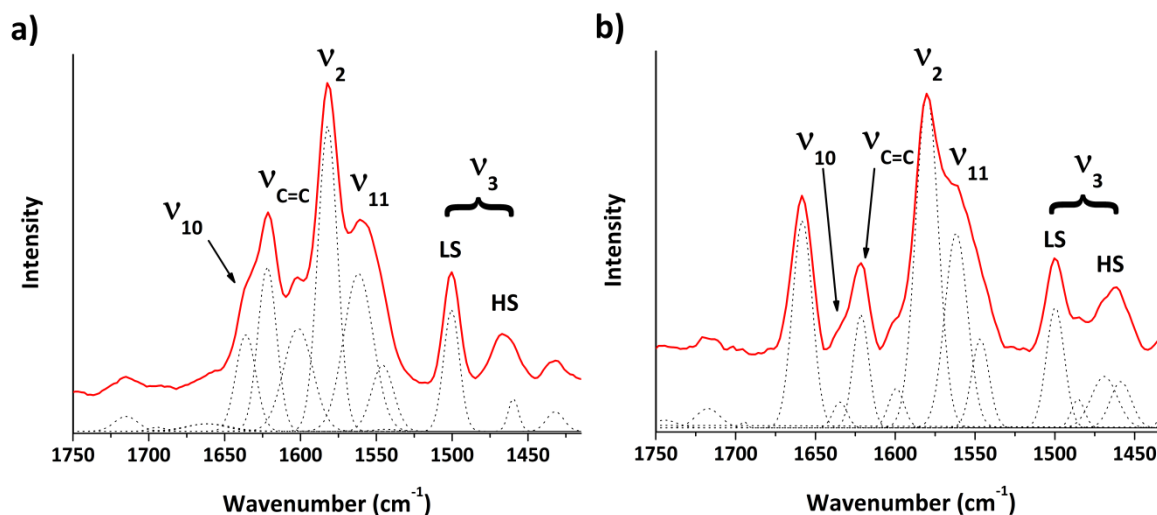


Figure 3.15. Sections ($1750 - 1325 \text{ cm}^{-1}$) of resonance Raman spectra of HDCYP102A3 ($\sim 200 \mu\text{M}$) a) substrate free, b) arachidonic acid ($\sim 1 \text{ mM}$) bound. Individual peaks from peak fitting are shown as dotted lines.

The spectra recorded (overview - Figure 3.14, expansions - Figure 3.15 & Figure 3.16) both show an intense signal at around 1370 cm^{-1} , the ν_4 band, which is an oxidation state marker indicative of ferric haem, and two smaller bands at around 1470 and

1500 cm^{-1} . These are the ν_3 bands, which are spin-state markers and are representative of the proportions of high-spin and low-spin ferric haem iron, respectively. There is a change in the relative intensities of these bands in the spectrum taken in the presence of arachidonic acid, which shows an increase in the proportion of high-spin haem in the sample[162]. This is consistent with substrate binding to the P450, displacing the water ligand from the haem and thus promoting a high-spin state. The signal at 1621 cm^{-1} is the $\nu_{\text{C}=\text{C}}$ feature, which seems to overlap with the ν_{10} signal, as seen before for BM3[163]. The ν_{10} occurs as a shoulder at 1635 cm^{-1} in both the substrate-free spectrum and the substrate-bound spectrum, although with reduced intensity in the latter. The signal at 1658 cm^{-1} in the substrate-bound spectrum (which does not appear to be present in the substrate-free spectrum) has not previously been assigned or reported. It is possible that it arises from an impurity in the arachidonic acid stock solution, although with Soret resonant excitation a novel feature arising from the haem is more likely.

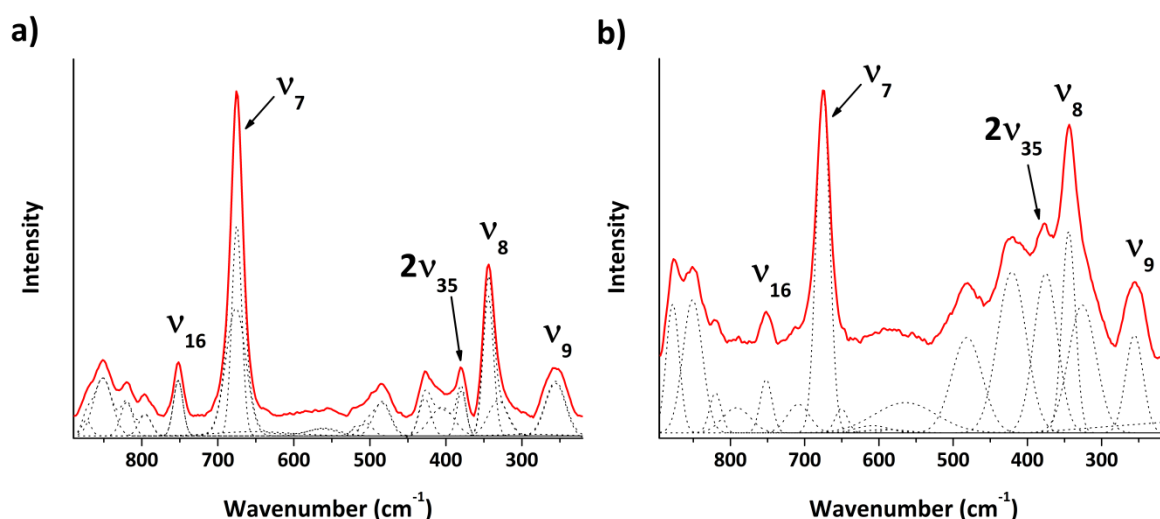


Figure 3.16. Sections (900 – 200 cm^{-1}) of resonance Raman spectra of HDCYP102A3 (~200 μM) a) substrate-free, b) arachidonic acid (~1 mM) bound. Individual peaks from peak fitting are shown as dotted lines.

Between 900 – 200 cm^{-1} are haem skeletal signals ν_{16} , ν_7 , $2\nu_{35}$, ν_8 and ν_9 [162] which remain largely at the same wavenumber in the substrate-bound spectrum. This may suggest that the haem itself is not particularly distorted by substrate binding. All signals discussed are presented in Table 3.5.

Signal	Wavenumber (cm ⁻¹)	
	Substrate	
	-	+
v₁₀	1635	1635
v_{C=C}	1621	1621
v₂	1582	1580
v₁₁	1560	1561
v₃	1500, 1467	1500, 1461
v₄	1371	1371
v₁₆	752	752
v₇	676	674
2v₃₅	379	377
v₈	344	344
v₉	256	256

Table 3.5. Selected vibrational bands present in the resonance Raman spectrum of HDCYP102A3 in the absence and presence of substrate.

All signals identified in the substrate-free spectrum are almost identical to those reported for BM3, excepting minor shifts in the v_7 and v_8 bands $674 \rightarrow 676 \text{ cm}^{-1}$ and $342.5 \rightarrow 344 \text{ cm}^{-1}$ respectively (BM3 \rightarrow CYP102A3)[163].

Overall, the similarity of the signals present suggests a high level of structural homology in the haem binding pockets of HDCYP102A3 and the haem domain of BM3. The increase, in the substrate bound spectrum, of the v_3 band at 1460 cm^{-1} , supports the results from substrate binding titrations (section 3.2.2.1) showing the development of a high-spin haem species upon substrate binding. Other haem resonance Raman bands are largely unchanged between substrate-free and substrate-bound spectra, indicating minimal structural changes in the haem environment (aside from coordination/spin state) upon substrate binding. The appearance of a signal at 1658 cm^{-1} on the substrate bound spectrum cannot be explained with reference to previous studies, either of other P450s or model systems. This would seem to indicate a previously unseen mode in resonance Raman spectra of P450s.

3.2.5. Redox Potentiometry

Redox potentiometry was carried out as described (section 2.4.13) to determine the redox potential of the haem of HDCYP102A3 in the presence and absence of substrate. Previous studies of the BM3 enzyme have revealed a substantial positive shift in haem potential on binding fatty acids, as is also the case for P450cam[111, 164].

Spectra (Figure 3.17) show a shift in Soret band absorbance from ~418 to 407 nm along with an overall decrease in Soret band intensity concomitant with reduction. Additionally, the α/β bands at 567 and 534 nm, respectively, are seen to merge into a single, broad band at around 550 nm. As the greatest change occurred at 418 nm, absorbance at this wavelength could be plotted against the potential relative to the standard hydrogen electrode to produce a potential curve, and by fitting using the Nernst equation a midpoint potential could be derived (Figure 3.20).

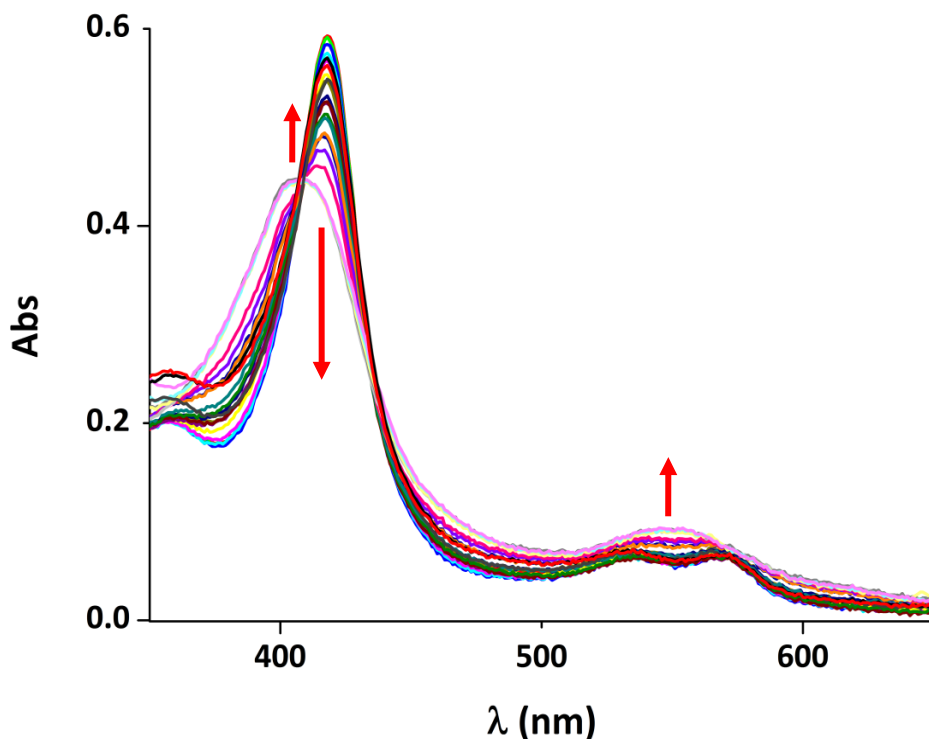


Figure 3.17. UV/Visible spectra of HDCYP102A3 ($\sim 6 \mu\text{M}$) under anaerobic conditions during reduction with sodium dithionite. A shift in the haem Soret band from ~ 418 to 407 nm can be observed, along with a decrease in absorbance. Also shown is the merging of the α and β features to form a single band at $\sim 550 \text{ nm}$ concomitant with reduction. The directions of changes in absorbance are indicated by red arrows.

The experiment was repeated in the presence of substrate NPG at ten times K_d ($74 \mu\text{M}$), assumed to be near-saturating conditions. In this case the Soret band shift was seen between different wavelengths: before addition of substrate the Soret absorbance was at 418 nm (as seen previously), but upon addition of saturating NPG it was seen to shift to 390 nm with approximately 25% decrease in absorbance and a significant shoulder at around 408 nm . These are features typical of the high spin form of the haem iron. Upon haem iron reduction the Soret band can then be seen to shift progressively to 410 nm (Figure 3.18). In this case changes in absorbance occurred across a range of wavelengths and so the approach used was to plot a difference spectrum (Figure 3.19) and then to summate the greatest changes in absorbance (at 437 and 389 nm) and to plot these values against the electrode potential to produce a curve (Figure 3.20). The data acquired in the presence of NPG were not of the same quality as those acquired in its absence. This may be due to turbidity, which can be seen in the spectra, due to substrate solubility limits.

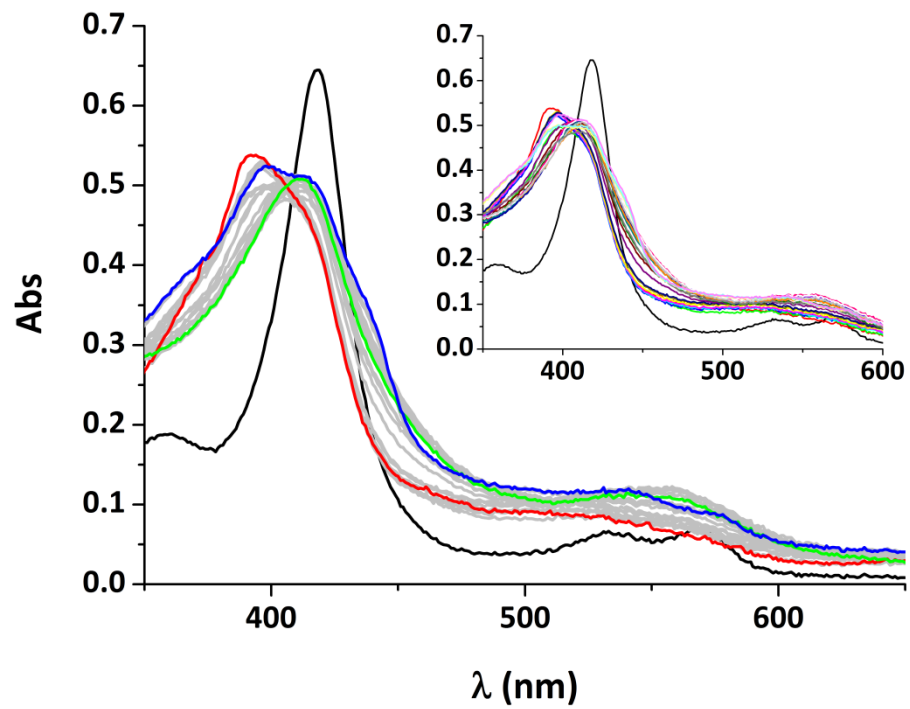


Figure 3.18. UV/Visible spectra of HDCYP102A3 under anaerobic conditions during sequential reduction with sodium dithionite in the presence of substrate *N*-palmitoyl glycine (NPG) to measure the $\text{Fe}^{2+}/\text{Fe}^{3+}$ redox potential. The experiment was carried out as described previously (2.4.10). The spectrum in black shows HDCYP102A3 ($\sim 7 \mu\text{M}$) with a Soret band at 418 nm. The spectrum in red shows the same after addition of NPG ($\sim 75 \mu\text{M}$) with the Soret band shifted to 397 nm and a residual shoulder at ~ 412 nm. All other spectra show the sequential reduction of the NPG-bound HDCYP102A3 with sodium dithionite. The spectral changes upon reduction are first a decrease in Soret absorbance and a shift to 412 nm (green spectrum), followed by the development of a peak at 398 nm (blue spectrum). In order to reveal the largest changes in absorbance, a difference spectrum was constructed (Figure 3.19). All spectra are shown in full colour in the inset panel.

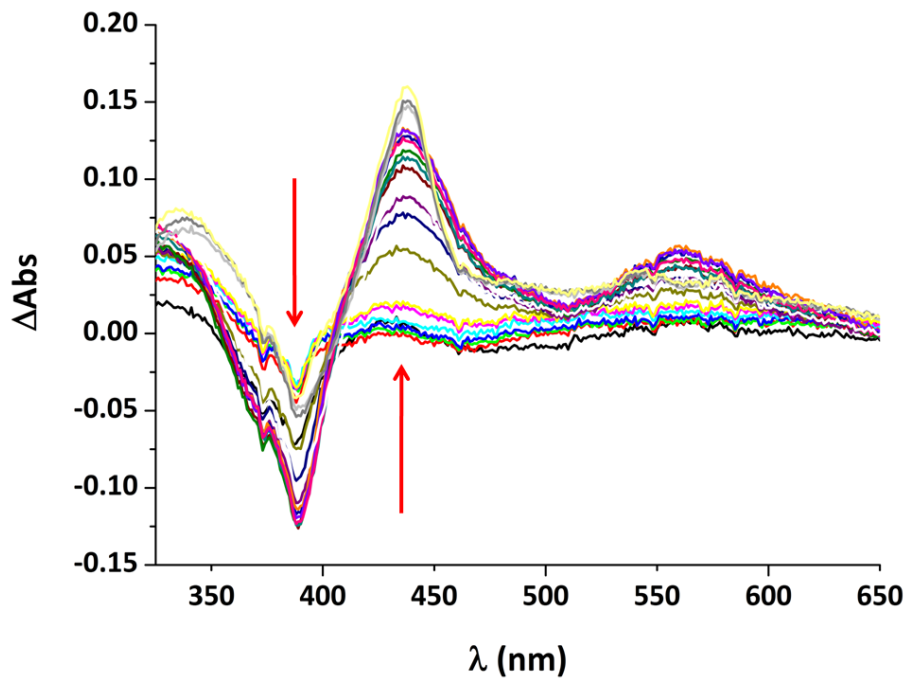


Figure 3.19. Difference spectra showing the wavelengths of maximum change during sequential reduction of HDCYP102A3 ($\sim 7 \mu\text{M}$) with sodium dithionite in the presence of substrate NPG ($\sim 75 \mu\text{M}$). A developing peak can be seen at 437 nm and a trough emerges at 389 nm. A further peak develops at ~ 560 nm as a consequence of the spectral fusion of features in the Q-band region.

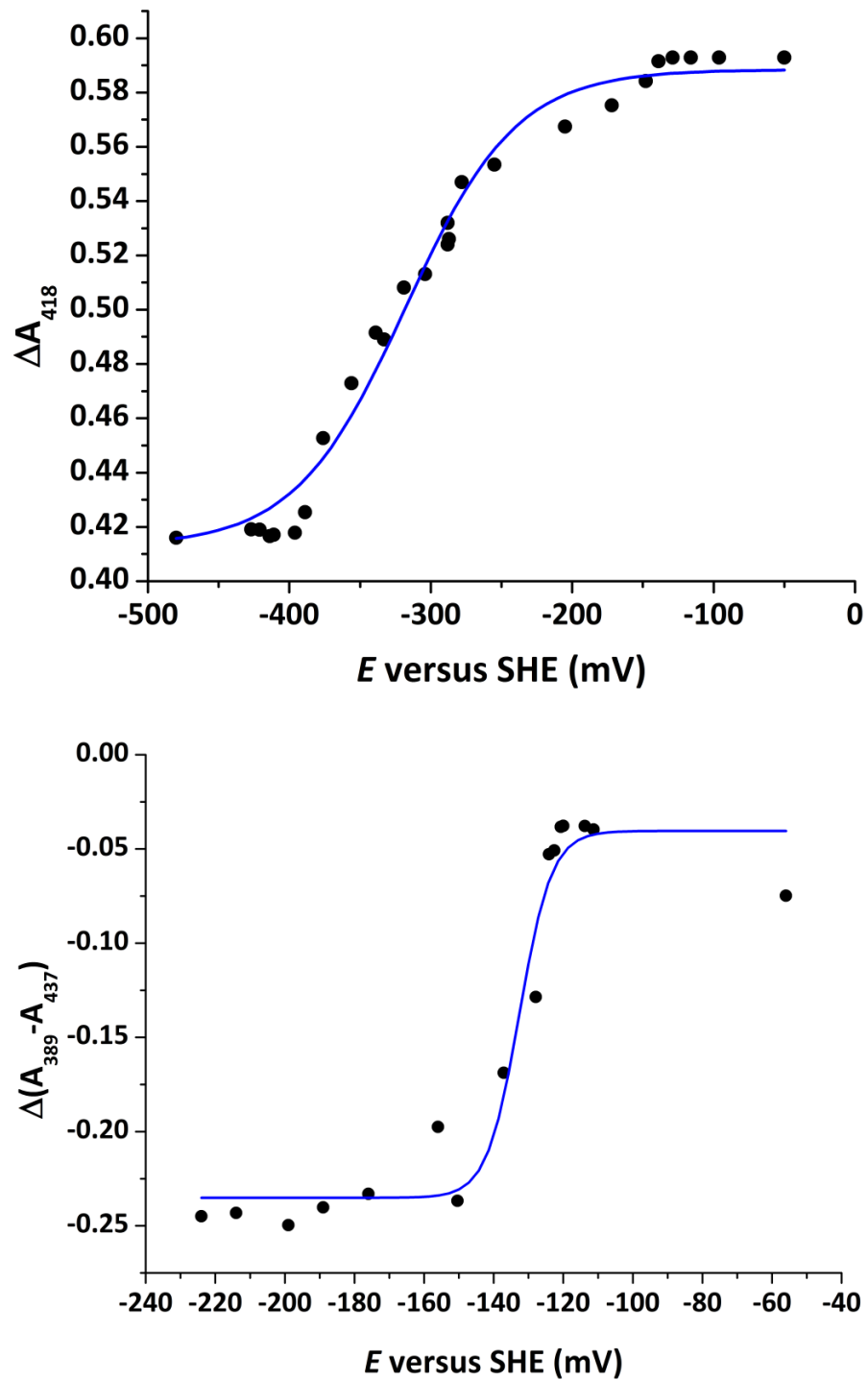


Figure 3.20. Redox potential curves for the haem domain of CYP102A3 as purified (*i.e.* low spin, top) and in the presence of substrate *N*-palmitoyl glycine (bottom). The x-axis is the potential against the standard hydrogen electrode and the y-axis is the change in absorbance at the Soret maximum 418 nm for substrate-free HDCYP102A3, and the difference in absorbance between 389 nm and 438 nm in the presence of NPG. Both are fitted using the Nernst equation for one electron transfer (blue lines) to determine midpoint potentials: -360 ± 8 mV (substrate free) and -133 ± 10 mV (NPG bound). The experiments were carried out as described previously (section 2.4.13)

It was shown, using the difference spectral data, that in the absence of substrate the haem iron $\text{Fe}^{3+}/\text{Fe}^{2+}$ standard potential was -360 ± 8 mV and in the presence of substrate NPG this was raised to -133 ± 10 mV. This represents a positive shift of at least 209 mV (within error limits) in the presence of substrate. It should be noted that there are fewer useful points in the titration in the presence of NPG, and thus that there is likely a greater error associated with the midpoint potential for the NPG-bound HDCYP102A3. Notwithstanding, there is clearly a substantial positive shift in haem iron potential in the NPG-bound P450.

The shift shows how the propensity for reduction of the haem increases upon substrate binding due to an increase in the proportion of haem iron in the high spin electron configuration. The magnitude of the shift measured is approximately double that measured for BM3 haem domain in the presence of substrates (100-130 mV)[111] and unprecedented for known P450s. The shift in midpoint potential may be exaggerated by data quality issues, mentioned previously, for the substrate-bound HDCYP102A3 data. It can nonetheless be said that a substantial shift in potential has occurred. The wavelength of the Soret band associated with the reduced species in each case was around 407-410 nm, which is consistent with retention of thiolate ligation of the haem upon reduction as seen in BM3 and P450cam[111, 165]. In some cases, notably CYP125A1 and CYP51B1 from *M. tuberculosis*, sodium dithionite reduction can induce protonation of the thiolate ligand to produce a thiol-ligated haem iron and with a Soret band typically shifting to a longer wavelength[37, 166]. This does not appear to be observed for HDCYP102A3 and confirms what was seen in the reduced CO-bound spectrum (Figure 3.4), where thiol-ligated haem would have appeared as a band at 420 nm[150]. The data provide clear evidence for the substrate-dependent elevation of haem iron potential in HDCYP102A3, which is likely required to enable electron transfer from the partner reductase enzyme in intact CYP102A3.

3.2.6. Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) was performed on HDCYP102A3 to investigate the unfolding profile of the enzyme and to establish, if possible, the

melting temperature(s) (T_m value or values) for the P450 and/or sub-domains thereof. Initially, large fluctuations in heat capacity measurements obscured any meaningful data but after several attempts with varying buffer conditions a somewhat cleaner unfolding profile was acquired.

At least two unfolding events seem to be visible from 40 – 50 °C, but unfortunately the quality of the data prevented accurate fitting to isolate the individual unfolding events. Additionally, at around 55 °C there was an extreme drop in heat capacity which is most likely due to wholesale aggregation of the sample. Though no specific T_m can be quantified, it can be said that HDCYP102A3 is at least stable at temperatures below ~40 °C and that there is at least one protein unfolding transition at around 50 °C. As the purity of the HDCYP102A3 stock has been established by various methods above, these results must be taken to show intrinsic properties of the protein as purified with respect to instability and aggregation. The temperatures at which these unfolding events and aggregation occur, fall below T_m values recorded for BM3 haem domain (58.7 °C and 64.9 °C)[112]. Despite this relative instability, it should be remembered that 50 °C is a temperature substantially higher than the physiologically relevant growth conditions for the host organism *Bacillus subtilis*, a bacterium typically found in soil.

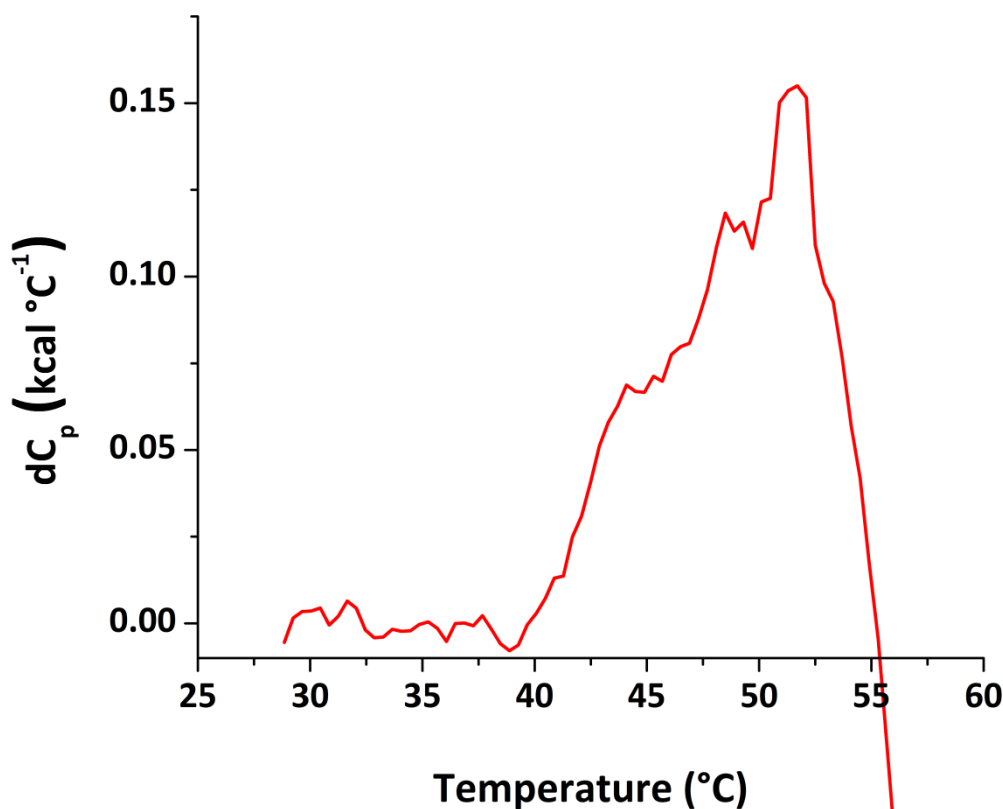


Figure 3.21. DSC trace for HDCYP102A3. The change in heat capacity (C_p) of a solution of HDCYP102A3 as temperature is increased is shown. Though the quality of the data prevents fitting to show individual peaks, there appear to be at least two unfolding events (at $\sim 45^\circ\text{C}$ and 50°C) that occur before the eventual, dramatic aggregation at $\sim 55^\circ\text{C}$.

3.2.7. Chemical Denaturation

In this series of experiments the effects on HDCYP102A3 of the denaturing agent guanidinium chloride (GdmCl) were examined. This is a chaotrope often used in studies of protein stability. Proteins incubated in 6 M GdmCl are usually fully denatured in terms of loss of their secondary and tertiary structure[167]. By using a range of concentrations of GdmCl and spectroscopic monitoring, the unfolding profile of a protein can be elucidated and the effect upon attributes such as cofactor binding or enzymatic activity explored. An example of the application of this technique to a P450 enzyme is the work of Lawson *et al.* on CYP107H1 (P450-BioI) from *B. subtilis*[168]. However, it has been shown that some secondary structure is retained in a variety of proteins under conditions of high GdmCl concentration, often indicating particularly compact structural elements[169]. Studies, mainly employing NMR and CD, have shown residual structure in 6 M GdmCl treated proteins. These include α -helices and β -sheets in staphylococcal nuclease[170] and hydrophobic clusters in chymotrypsin inhibitor 2[171].

GdmCl denaturation was employed here to provide additional enzyme stability data in light of the issues with respect to the DSC data discussed above. Identical samples of the HDCYP102A3 protein ($\sim 1.8 \mu\text{M}$) were incubated with a range of concentrations of GdmCl for 10 minutes and then the denaturation monitored by UV/Visible spectroscopy measurements.

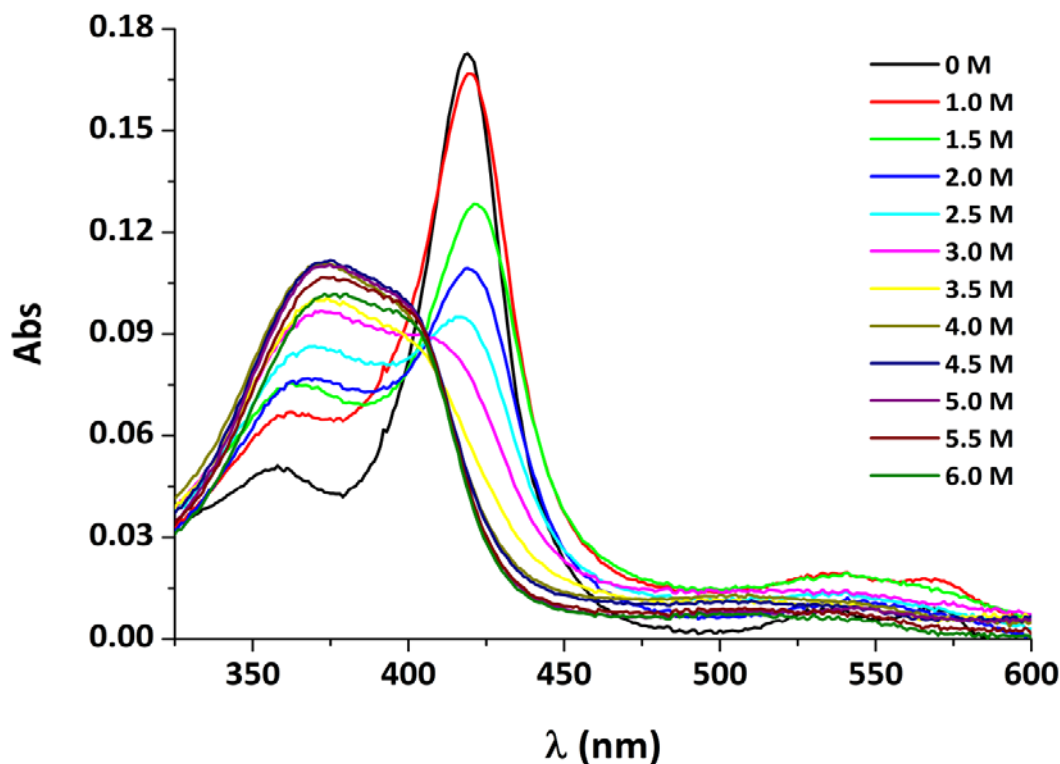


Figure 3.22. UV/Visible Spectra showing changes in absorbance of HDCYP102A3 ($\sim 1.8 \mu\text{M}$) following incubations with different concentrations of the denaturant guanidinium chloride. The legend shows the molar concentration of denaturant corresponding to each spectrum.

The UV/Visible spectra recorded during the experiment are shown in Figure 3.22. These reveal a progressive diminution of the Soret peak at $\sim 418 \text{ nm}$ and the concomitant development of a broad absorption feature between $\sim 350\text{-}400 \text{ nm}$. This likely reflects the progressive dissociation of the P450 haem from its proximal thiolate ligation, leading to the displacement of the haem from its binding site and ultimately from the protein matrix itself. Difference spectra were produced to highlight the wavelengths of maximum change. These were produced by subtracting the spectra of the sample in the absence of GdmCl (0 M in the figure) from those recorded at each of the other denaturant concentrations. The difference spectra are shown in Figure 3.23.

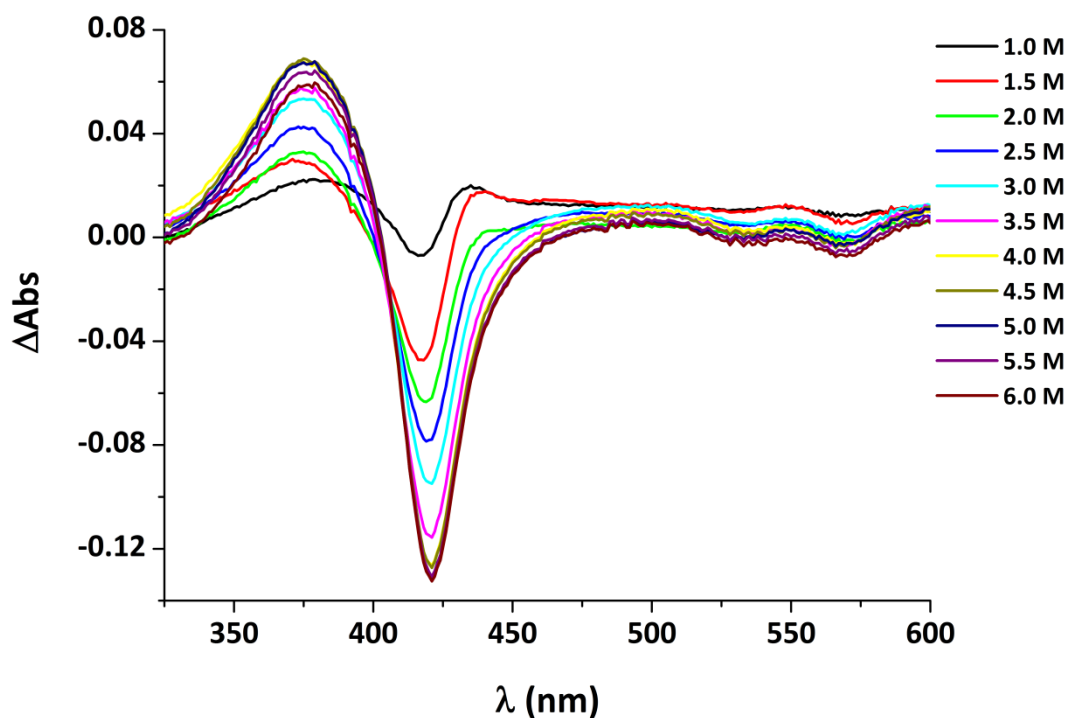


Figure 3.23. Difference spectra showing the change in absorbance of HDCYP102A3 (~1.8 μM) on incubation with varying concentrations of the chaotropic denaturing agent guanidinium chloride (GdmCl). The legend shows the molar concentration of denaturant corresponding to each spectrum. As the concentration of GdmCl is increased the haem Soret band absorbance is seen to shift with a trough forming at 421 nm and a peak growing at 375 nm, indicating the displacement of the haem thiolate ligand, and release of the haem group from its binding site, and then from the protein itself upon unfolding of the protein. An isosbestic point is visible at 405 nm. The α/β features between 500 – 600 nm are also gradually lost as the concentration of denaturant is increased and haem is displaced.

From the difference spectra it is clear that the haem Soret signal shifts from 421 nm, normal for P450-bound haem, to 375 nm consistent with free haem[172]. This suggests that, as the concentration of GdmCl is increased, increasing proportions of the protein are denatured and lose their haem. Further, with single Soret trough and peak features at 421 nm and 375 nm, respectively, it appears that two major species predominate (low-spin thiolate-coordinated haem and free haem that has lost its protein ligation) during the denaturant titration.

In order to show a progress curve for denaturation and to quantify the loss of haem and establish a midpoint value of GdmCl concentration relating to the stability of the HDCYP102A3, changes in absorbance values at 421 nm were plotted against the concentration of GdmCl (Figure 3.24). These data were fitted using a denaturation equation (Equation 3.1) which provides a value of the Gibbs free energy of unfolding (ΔG^0) of $7.88 \pm 0.71 \text{ kJmol}^{-1}$.

$$y = y_D + \frac{(y_D - y_N)}{1 + e^{\left(\frac{\Delta G^0 - mx}{-RT}\right)}}$$

Equation 3.1. Denaturation equation used for the fitting of chemical denaturation data. Where y_N and y_D are the values of y in native and denatured states respectively. ΔG^0 is the standard Gibbs free energy of unfolding. mx is the dependency from the concentration of denaturant and R and T are the ideal gas constant ($8.314 \text{ J K}^{-1} \text{ mol}^{-1}$) and temperature in Kelvin respectively.[167]

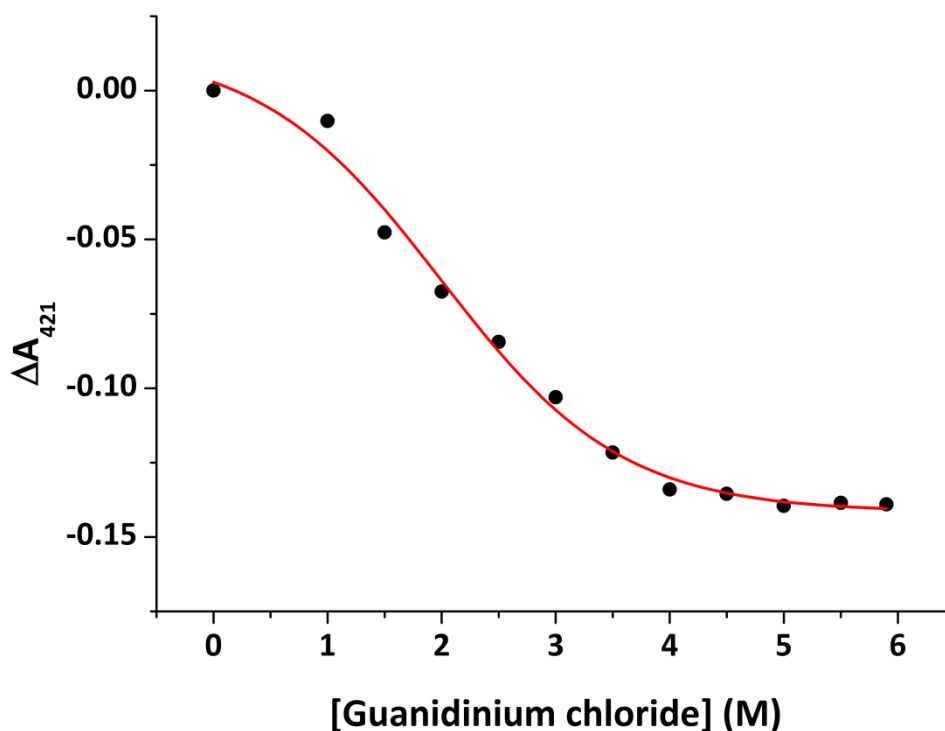


Figure 3.24. Denaturation curve for HDCYP102A3 ($\sim 2 \mu\text{M}$) showing the change in haem Soret absorbance at varying concentrations of the chaotropic denaturing agent guanidinium chloride. In this way the stability of the haem group during unfolding of the protein is shown. The midpoint of the curve, at which the unfolding event is 50% complete, is $2.1 \pm 0.2 \text{ M GdmCl}$. The value of the Gibbs free energy for the transition is $7.88 \pm 0.71 \text{ kJmol}^{-1}$.

Excepting a slight increase in turbidity at higher concentrations of GdmCl, the denaturation profile does not show the aggregation seen in the DSC experiment. This is almost certainly a consequence of the solubilising effects of the denaturant, which overcomes any major tendencies for protein association and precipitation. Additionally, recent literature has suggested that during guanidinium chloride denaturation some secondary structural elements can be retained in certain proteins and that these include hydrophobic clusters[169]. This could prevent aggregation

due to intermolecular hydrophobic interactions between protein molecules and might be unlikely to persist under conditions of thermal denaturation.

In experiments with BM3 it was found that at concentrations above 0.6 M GdmCl, catalytic activity was diminished to less than 5% of the value in the absence of GdmCl. It was also suggested from this work that the loss of catalytic activity might be mainly attributable to unfolding in the reductase domain, and specifically to the dissociation of the FMN, which is the more weakly bound of the flavin cofactors[112]. If the results for HDCYP102A3 are representative of the case in BM3 then these results also support this hypothesis, with the haem domain not showing significant alterations in the haem-thiolate spectrum or loss of haem at 1 M guanidinium chloride concentration.

3.2.8. Light Scattering

Multi-angle light scattering (MALS) can be used to establish the monodispersity of a protein stock. It can detect oligomers and the proportion in which they are present. By measuring hydrodynamic radius it can also hint at how tightly the protein is folded – *i.e.* how compact its structure is[173].

MALS was carried out on the purified stock of HDCY102A3 prior to crystallographic trials to ensure that successful crystallisation was not precluded by a polydisperse or otherwise heterogeneous sample. The MALS experiment involved loading the HDCYP102A3 sample onto a size exclusion column. The flow-through from this column was then monitored for light scattering and absorbance at 280 nm. The trace from these detectors (Figure 3.25) shows a single peak eluting from the column after ~14 mL. This shows that there was a single population of protein species present in the HDCYP102A3 stock.

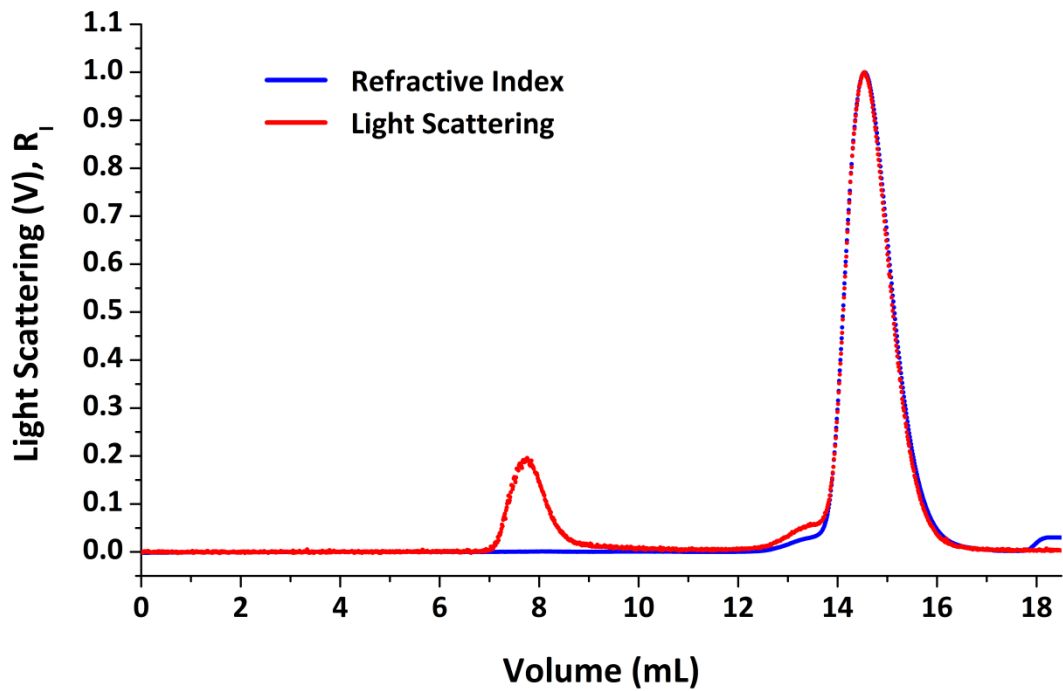


Figure 3.25. Trace from a Multi angle light scattering (MALS) run of a sample of HDCYP102A3. The x-axis shows the elution volume from the size exclusion column and the y-axis shows a measurement from two detectors. Light scattering is indicated as a voltage and refractive index (R_i) is a relative scale. The single peak occurring between $\sim 14 - 16$ mL shows one species of HDCYP102A3. The peak between $\sim 7 - 8.5$ mL is known as the void and is expected in this experiment.

From the data acquired, values for the hydrodynamic radius (R_h) and observed molecular mass (M_r) could be derived by a method known as Zimm fitting[136, 173]. All fitting was handled by the software (Astra by Wyatt Technology Corporation) associated with the apparatus (a Dawn EOS instrument by Wyatt Technology Corporation). When the values of R_h and M_r are plotted against elution volume (Figure 3.26) a steady value of each parameter over the duration of the peak is seen. This indicates that the HDCYP102A3 stock is highly monodisperse, with a single globular protein species. Zimm fitting provided values of the average hydrodynamic radius at 2.8 nm and the molecular mass at 55.4 kDa indicating a monomer. The slight difference between the value for the molecular mass determined by MALS and those observed for HDCYP102A3 on a SDS PAGE gel and predicted from the amino acid sequence (52 kDa) occurs mainly as mass is inferred in MALS from the hydrodynamic radius of the sample. A monodisperse, homogenous sample such as this is more likely to produce useful crystals than one containing multiple species[174].

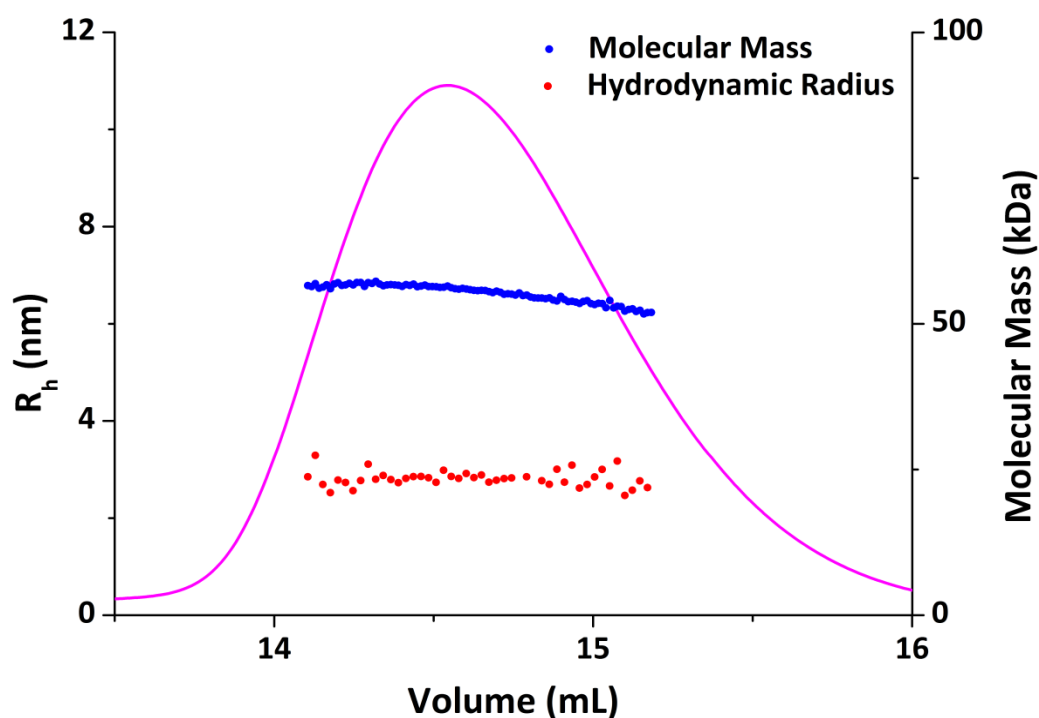


Figure 3.26. A plot of values, derived from Zimm fitting, of the hydrodynamic radius (R_h) and molecular mass against elution volume for a multi angle light scattering (MALS) run of a sample of HDCYP102A3. In magenta is the peak in the original trace for comparison. The molecular mass varies from ~56-52 kDa which matches closely with that predicted from the amino acid sequence of HDCYP102A3 (55.2 kDa). The peak hydrodynamic radius is 2.8 nm.

3.2.9. Crystallographic Trials

With the monodispersity of the HDCYP102A3 sample verified by MALS, initial crystal trials were set up on 96-well plates (as described in section 2.4.10). A 12 mg mL⁻¹ solution of HDCYP102A3 was used with a number of commercially available screens, namely: Classics, Classics II, Classics L, Ammonium sulfate suite, Cations suite, Anions suite, PEGs suite, JCSG+ suite, Mb Class I, Mb Class II and ProComplex (all supplied by QIAGEN).

Hexagonal HDCYP102A3 crystals were obtained in several wells, the most promising of which (in terms of crystal viability) contained 0.8 M sodium citrate at pH 6.5 in the mother liquor.

Further crystallisation trials were carried out using substrate-bound HDCYP102A3 with a similar range of screening conditions, using a 10 mg mL⁻¹ HDCYP102A3 solution incubated with substrate NPG (14 μ M). Crystals of sufficient size, for x-ray diffraction experiments, were formed in the nanolitre scale trials using 0.1 M Tris, pH 8.5, 25 % (v/v) *t*-butanol as mother liquor. The substrate bound crystals were hexagonal, like those produced from the substrate-free stock.

In both cases, crystals appeared to be of good quality in terms of size and regularity of shape (see Figure 3.27), and the protein consistently crystallised in the hexagonal form. While the colouration of the crystals is consistent with haem containing HDCYP102A3, it was noted that other P450 enzyme crystals are of a distinctly deeper red colour.

Unfortunately, neither crystal form diffracted beyond $\sim 15 \text{ \AA}$, despite trialling several crystals using a range of flash-cooling conditions and high-intensity synchrotron radiation. It is possible these crystals have an unusually high solvent content (a possible explanation for the relatively pale colour of the crystals) which has led to the apparent high degree of disorder that limits the diffraction. Despite further trials, no other crystal forms could be obtained.

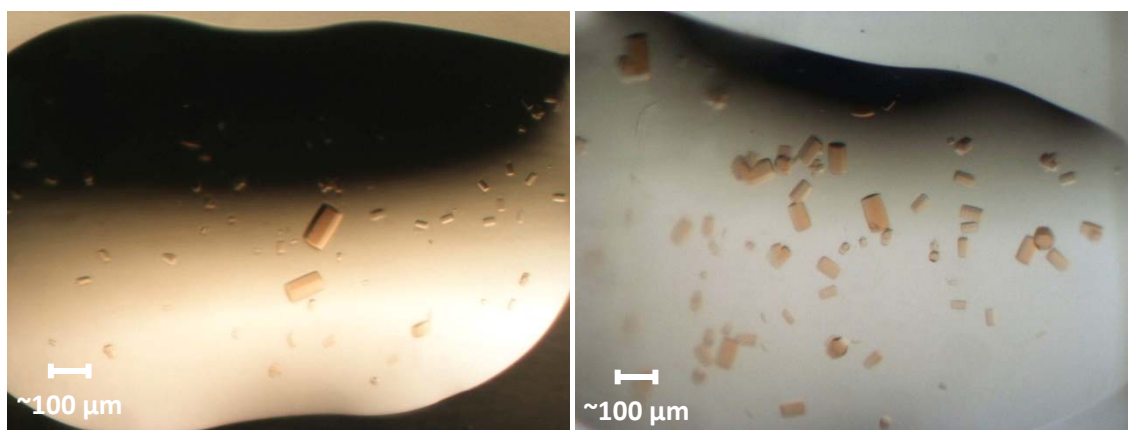


Figure 3.27. Images of the HDCYP102A3 crystals obtained, isolated in the absence (left) and presence (right) of substrate NPG (14 μM). The conditions under which crystals formed were 0.8 M citrate at pH 6.5, and 0.1 M Tris at pH 8.5 with 25 % (v/v) *t*-butanol respectively.

3.2.10. Homology Modelling

In the absence of a crystal structure of HDCYP102A3 from either the substrate-free or substrate-bound trials, homology modelling was undertaken. This technique aims to build a model of a protein structure based on an existing (template) structure of a homologous protein. In the case of HDCYP102A3 an obvious choice is BM3 for which, as stated before (section 3.1), many structures exist. The model was created using Swiss-Model, an online service by the Swiss Institute of Bioinformatics, which handles fully automated modelling of proteins from their amino acid structures[175-177]. Being fully automated, the service also selected an appropriate template structure from the RCSB Protein Data Bank. The template structure chosen was the structure of the BM3 haem domain taken from the crystal structure of the linked BM3 haem and FMN-binding domains, which generated a non-stoichiometric and proteolytically separated structure of two haem domains for each FMN domain[137] (RCSB PDB id: 1BVY).

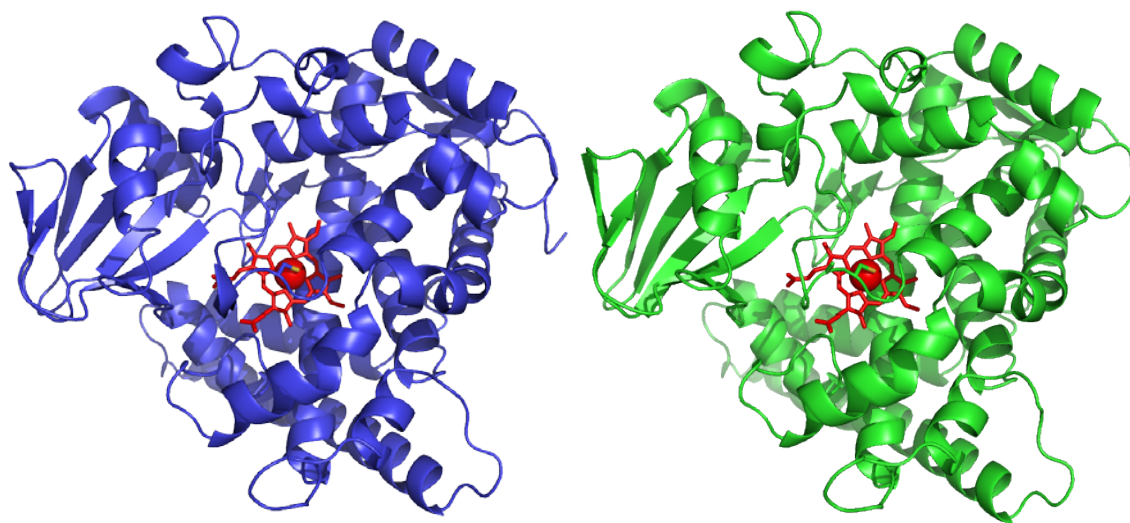


Figure 3.28. Side by side comparison of the template structure of BM3 haem domain (blue) and the model of HDCYP102A3 based upon it (green) shown as cartoons. In each image the haem (red) and the ligating cysteine are shown. Aside from the orientation and position of the two β -sheet structures towards the left of each structure, there are no major differences in the physical structure.

The model reflects the high degree of homology between the two proteins, showing a very similar overall shape to the template structure (see

Figure 3.28). This should be expected as, due to the nature of homology modelling, the HDCYP102A3 amino acid sequence has effectively been arranged into a BM3 conformation before undergoing energy minimisation. The resulting model may or may not reflect the actual conformation of HDCYP102A3 and so some caution is required with interpretation.

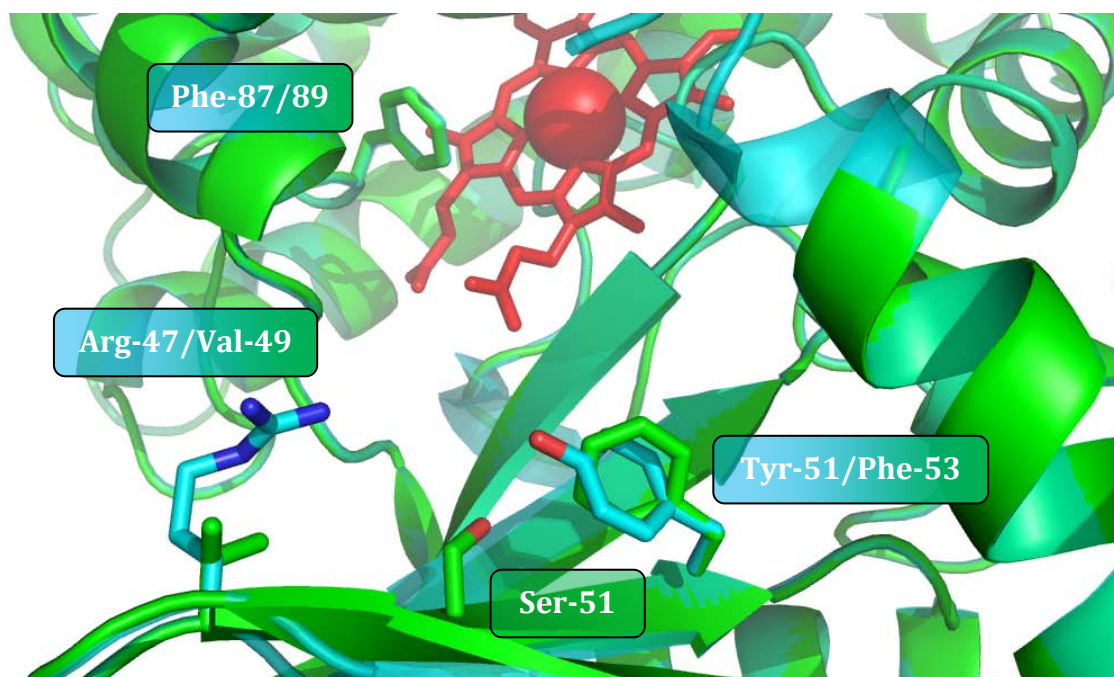


Figure 3.29. Close-up image of the homology model of HDCYP102A3 (green) with the template structure (HD-BM3) overlaid (transparent cyan), and with both shown as a cartoon. Shown as sticks: In the foreground are Arg-47 and Tyr-51 in BM3 and Ser-51 in HDCYP102A3. Towards the background are the haem group (red) and the phenyl rings of Phe-87/Phe-89 (BM3/HDCYP102A3).

An obvious difference between the two structures is the absence in HDCYP02A3 of equivalents for Arg-47 and Tyr-51, residues known to form hydrogen bonds with the carboxylate group of fatty acid substrates in BM3. Occupying the same space as these residues in the HDCYP102A3 model are Val-49 and Phe-53 respectively, suggesting that the same hydrogen bond anchoring does not occur in CYP102A3. The only residue in the same area of the model that could have a similar role is Ser-51. A view along the substrate binding cavity towards the haem showing these residues is presented in Figure 3.29.

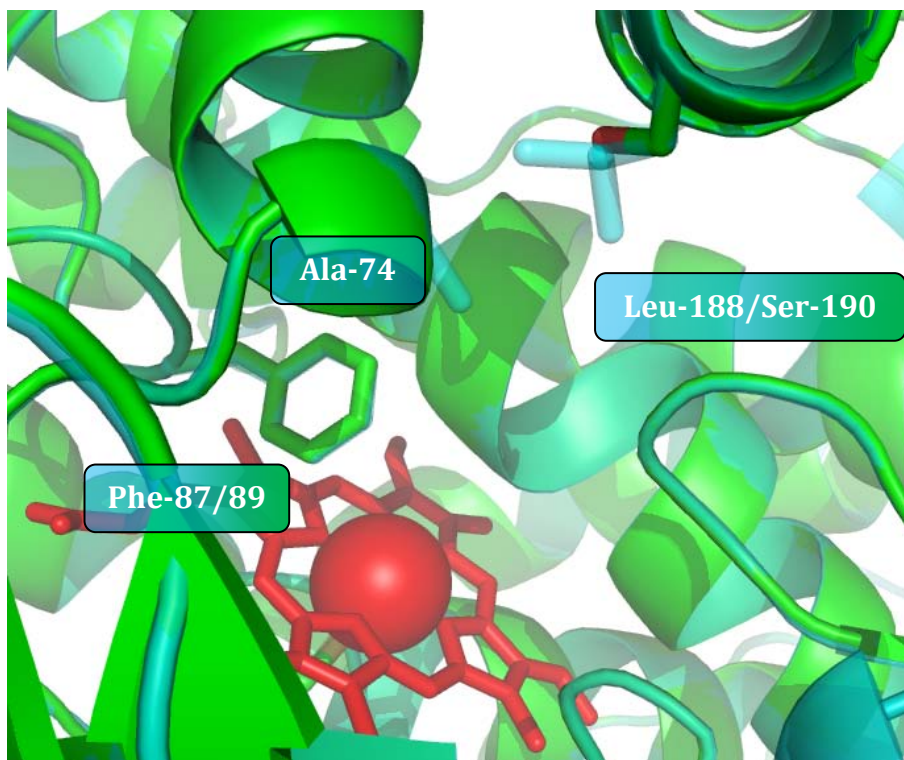


Figure 3.30 Close-up image of the homology model of HDCYP102A3 (green) with the template structure (HD-BM3) overlaid (transparent cyan), and with both shown as a cartoon. Shown as sticks: In the upper half of the image, Ala-74 and Leu-188 in BM3 and Ser-190 in HDCYP102A3. In the lower half of the image are the haem group (red) and the phenyl rings of Phe-87/Phe-89 (BM3/HDCYP102A3).

With the important caveats, mentioned above, acknowledged, this homology model does demonstrate what the effects of the key differences in amino acid sequence, between the haem domains of BM3 and CYP102A3, would be. The absence of the BM3 carboxylate-anchoring residues in CYP102A3 leaves no comparable hydrogen bonding residues present to fulfil this role. A fatty acid substrate could, therefore, be imagined to bind further into the substrate binding cavity and enjoy better access to the haem for its higher sub-terminal carbon centres.

3.3. Summary & Discussion

The work presented in this chapter covers the biochemical and biophysical analysis of HDCYP102A3, a homologue of the well characterised model enzyme BM3 (CYP102A1). A key aim in this work was the structural characterisation of HDCYP102A3 by x-ray crystallography. Though this aim was unsuccessful, due to insufficient diffraction, a large amount of useful characterisation of the HDCYP102A3 protein was carried out.

A homology model of HDCYP102A3, based on its amino acid sequence and its closest structurally characterised homologue BM3, showed a very similar overall structure. It is, however, difficult to be certain of the similarity from this result, as the technique is likely to produce a model similar to the template structure.

EPR and resonance Raman experiments provided some structural information regarding the haem centre of HDCYP102A3. The EPR spectrum recorded displayed the characteristic signal of a low-spin, hexacoordinate, cysteine ligated haem. The g-values of the signals in the EPR spectrum of HDCYP102A3 are almost identical to those for the haem domain of BM3. This confirms a high level of structural similarity between the two. Resonance Raman spectra showed a strong agreement between most signals for HDCYP102A3 and those found in spectra of BM3, presenting a picture of a haem environment very similar to that found in BM3. These two results both support the view that the homology model is a fair representation (at least of the immediate environment of the haem) of the structure of HDCYP102A3.

Resonance Raman spectra were also recorded with and without bound substrate, and these spectra showed the characteristic shift towards a high-spin haem iron upon binding of fatty acid substrates. Binding titrations were carried out to quantify the binding affinities of a selection of substrates and inhibitor molecules. These experiments revealed that the substrate and inhibitor binding affinities of HDCYP102A3 are similar to those recorded previously for the full length fusion enzyme, with some notable differences. A pattern of higher binding affinity for longer substrate chain lengths was shown, that was not present in the values for the full length enzyme. Additionally, the overall preference for branched-chain fatty acid substrates, shown for the full length enzyme[79], was not evident in the results for the isolated haem domain. Though no binding interaction was observed for palmitic acid while 15-methylpalmitic acid did bind, this pattern was not borne out by results

for other branched chain substrates. In the case of 12- and 13-methylmyristic acids, K_d values were greater than that for myristic acid (the equivalent straight chain substrate). These results point to a possible influence upon substrate binding of the presence of the reductase domain and the likely dimeric nature of the full length enzyme. Binding of NPG, a derivatised fatty acid shown to bind tightly to BM3, was confirmed for HDCYP102A3, albeit with ~20 fold lower affinity (based on K_d values)[148].

The binding of inhibitors was also shown to differ, with twofold tighter binding in HDCYP102A3 (compared to the full length enzyme) in the case of imidazole and 1-phenylimidazole. Conversely, HDCYP102A3 was shown to have a tenfold reduced affinity for 4-phenylimidazole. Additionally, an inhibitor producing no binding interaction with BM3, 2-phenylimidazole, was shown to bind to HDCYP102A3 likely in a similar conformation as it has been found to bind in P450cam. All of these variations suggest that, despite the structural similarities between CYP102A3 and BM3 haem domains, the binding interactions between ligands and amino acid side chains in the substrate binding cavity (possibly including Phe-89 in CYP102A3, Phe-87 in BM3) are significantly altered.

Redox potentiometry experiments showed an increase in the haem Fe^{2+}/Fe^{3+} standard potential in the presence of bound substrate. While a positive shift in this redox potential is expected for a substrate-bound P450 and has been demonstrated in many cases, the magnitude of the increase is unprecedented (>200 mV) and warrants some caution. However, structural influences on the haem (e.g. planarity of the macrocycle and hydrogen binding interactions with its substituent groups) could compound the effects mediated by the high-spin transition in the substrate-bound form, and underlie the larger than expected change in potential observed in substrate-bound HDCYP102A3.

Studies of the stability of HDCYP102A3 to denaturing conditions were carried out by DSC and chemical denaturation. The DSC data acquired, while suffering from background noise, show what appear to be two unfolding events with T_m 45-50 °C followed by massive aggregation of the HDCYP102A3 protein. These two unfolding events could well correspond to the separate unfolding of the α and β domains of the protein. This experiment also showed that the protein is stable at temperatures well above what the host species (a soil bacterium) would normally experience.

Chemical denaturation with varying concentrations of GdmCl was used to measure loss of haem during unfolding. It was shown that at concentrations of GdmCl that were highly deleterious to the catalytic activity of BM3 (1-2 M) the effect on the bound haem of HDCYP102A3 was low. This result supports the suggestion that the loss of activity at < 1 M GdmCl in BM3 is due mainly to reductase domain unfolding and FMN dissociation. The GdmCl experiment did not show the same protein aggregation as seen in the DSC experiment. This may be due to the solubilising effect of the denaturant or due to retained secondary structure (with the former being more probable). It should also be noted that the isolated haem domain of CYP102A3 may have a lower stability than the full length enzyme, especially if the latter exists as a dimer which may provide further stability. The HDCYP102A3 protein appeared to exist as two distinct species during this experiment: a low-spin, thiolate ligated haem-containing, intact protein and a denatured protein having lost its haem. No intermediate, partially denatured, spectral species were observed from optical studies. Further studies using far UV, CD or aromatic amino acid fluorescence could provide such information, although one might expect the dissociation/loss of ligation of the haem cofactor to occur at a GdmCl concentration lower or similar to those required for gross conformational perturbations of the protein.

To address concerns about aggregation of HDCYP102A3, especially with regard to planned crystallographic experiments, a MALS experiment was carried out to ascertain the monodispersity of the HDCYP102A3 protein in solution. This experiment showed that the protein was indeed highly monodisperse with negligible levels of aggregation. Encouraged by this result, crystallographic trials were carried out, leading to the generation of two crystal isoforms: one with bound substrate (NPG) and one without. Unfortunately neither of these crystal samples produced sufficient diffraction for structural characterisation. After these discouraging results, studies of HDCYP102A3 were concluded and work commenced on the CYP116B1 fusion system (see Chapters 4 & 5) with the aim of generating data on the structural and biochemical characteristics of this novel class of P450-redox partner fusion enzyme.

Going on from the work on HDCYP102A3 presented in this chapter, further efforts towards structural characterisation could be pursued. The best approach would be to attempt to produce a different crystal isoform, by employing a range of ligands

(substrates or inhibitors) likely to produce conformational change in the protein (or to stabilise and alternative P450 conformer or conformers). Based on the structural similarity highlighted in this work and the homology model, mutagenesis experiments could be undertaken in a similar fashion to those carried out on BM3 to probe various aspects of the enzyme. These include analysis of the effects of mutating key residues on the substrate selectivity, regioselectivity of substrate oxidation and the thermal stability of CYP102A3. The insight gained in these experiments could form part of an overall approach aimed at combining desirable features from all of the CYP102A subfamily enzymes to produce a range of novel biocatalysts for specific applications such as those discussed in the literature on the engineering of BM3 (see section 1.1.10).

Characterisation of Novel Cytochrome P450 Fusion Systems

Chapter 4

CYP116B1 Protein Studies

4. CYP116B1 Protein Studies

4.1. General Information

4.1.1. Survey of P450 Fusions

Following the conclusion of work on CYP102A3, another P450 fusion system of interest was chosen for study. As stated in the introduction, BM3, with its C-terminal CPR-like domain, heralded the discovery of P450 enzymes with fused reductase systems (see section 1.1.9). In the years following this, a number of direct homologues have been discovered and, interestingly, several other P450 fusions with different types of reductase have been identified. Despite these discoveries, it seems that much remains to be elucidated in the variety of reductase systems that can be found fused to P450 enzymes.

The conserved domain architecture retrieval tool (CDART) performs a search based upon a protein amino acid sequence to find genes encoding proteins with similar arrangements of functional domains[178, 179]. A CDART search using the amino acid sequence of BM3 reveals several known classes of P450 fusion enzymes and many others which constitute a P450 fused to a known protein domain. In the tables presented and discussed below are presented a selection of the results from the CDART search. Those fusion P450s whose descriptions are in bold are those that have currently been characterised and those in italic are predicted from genetic data from sequencing projects of the organisms in question.

4.1.1.1. *P450-Redox Partner Fusions*

Of all the types of P450 fusion proteins described in this section, the redox partner fusions are the most widely studied. Practically all of the characterised P450 fusion proteins fall into this category. These characterised examples and a few potentially novel examples are presented in Table 4.1.

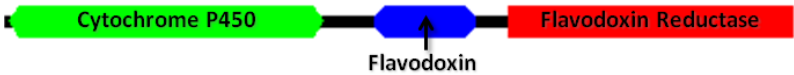

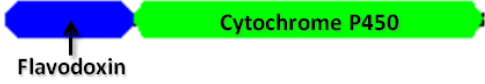
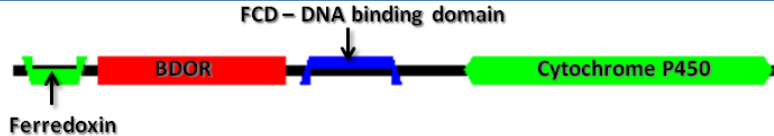

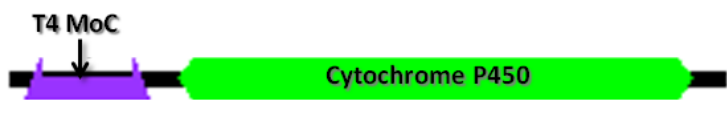
Fusion type	Layout	Number of sequences
CPR-like reductase domain		153
PDOR-like domain		56
Flavodoxin-like domain		16
<i>Benzoate dioxygenase reductase</i>		3
<i>Rossmann fold-type NAD(P)H binding domain</i>		5
<i>Toluene monooxygenase ferredoxin</i>		2

Table 4.1. Selected results from a CDART search using the amino acid sequence of P450 BM3 showing domain layouts of established (**bold**) and potential (*italic*) P450 fusions. Graphical representations of the domain layout are annotated from the CDART tool[178, 179]. PDOR – Phthalate dioxygenase reductase, BDOR – benzoate dioxygenase reductase, CAD – cinnamyl alcohol dehydrogenase and FCD – FadR C-terminal domain.

The first fusion P450 in Table 4.1. is the model for BM3 and most other CYP102 and related enzymes such as P450-foxy (CYP505A1)[75]. These enzymes contain a P450 haem domain fused to a CPR-like reductase domain comprising NAD(P)H flavodoxin reductase and flavodoxin-like subdomains. Discussion of this class of fusion P450 is covered in Chapter 3.

The second fusion comprises a P450 haem domain fused to a reductase domain resembling phthalate dioxygenase reductase (PDOR) and consisting of NAD(P)H-ferredoxin reductase-like and ferredoxin-like subdomains. The 56 sequences listed for this class of P450 fusion show a high degree of amino acid sequence identity (38-78 %). A phylogenetic tree of the sequences is shown in Figure 4.1. All of the P450s in this class appear to belong to, or to be homologous to, the CYP116 family of P450 enzymes.

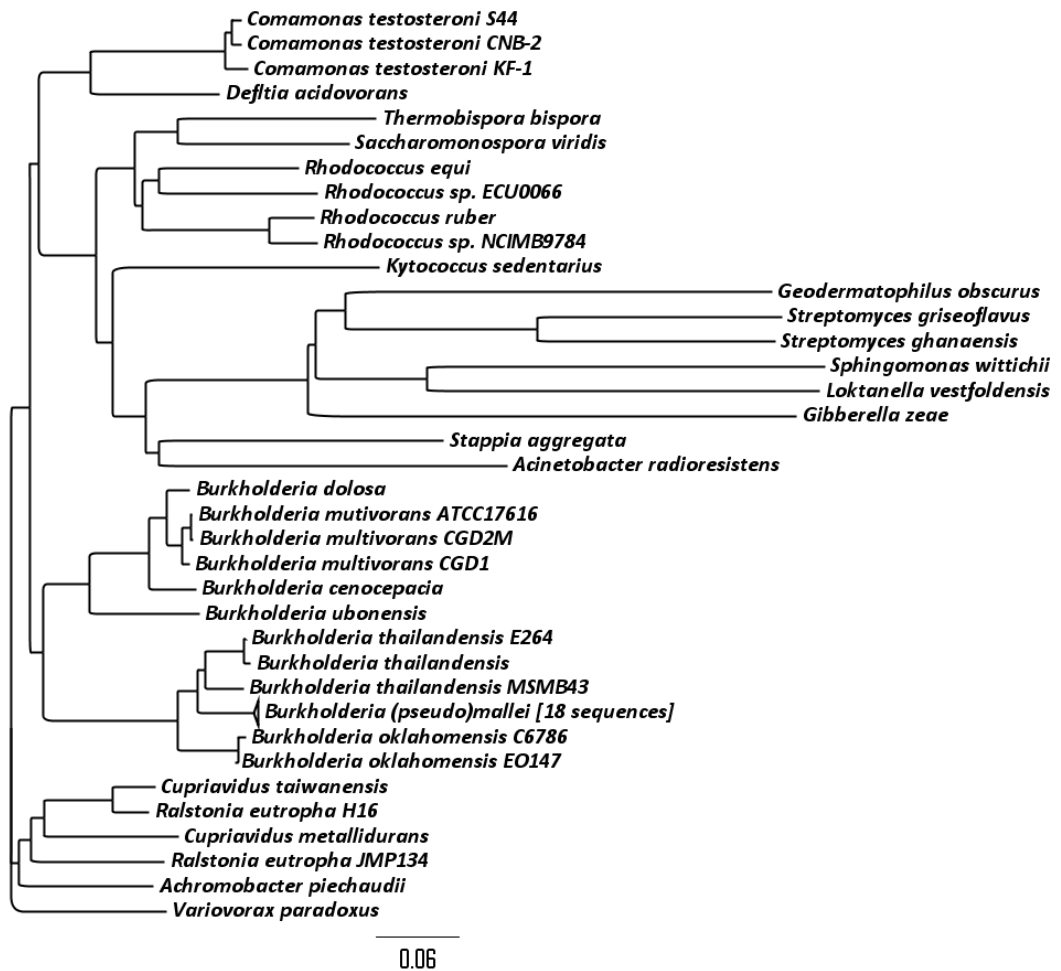


Figure 4.1. Phylogenetic tree of amino acid sequences listed as P450-phthalate dioxygenase reductase-like fusions in a CDART search using the amino acid sequence of P450 BM3. Each sequence is represented by the name of the host organism while 18 sequences from organisms of *Burkholderia mallei* and *pseudomallei* strains are collapsed into a single entry for ease of viewing and as these are highly similar sequences.

The third type of P450 fusion enzyme in Table 4.1 is a P450 fused to an N terminal flavodoxin (as opposed to the C-terminal fusion in CYP102 type enzymes). This class of fusion is not catalytically self sufficient in the way that the first two are, as it has no ability to abstract electrons from NAD(P)H. It therefore requires a redox partner protein to fulfil this role and to complete its catalytic cycle. Though 16 sequences are returned by CDART for this class of fusion, only two of the sequences appear to contain a full P450 domain. The remaining sequences do not incorporate a haem binding motif and so cannot possibly represent full P450 fusions. This may be the true nature of the proteins or due to incomplete sequencing of the relevant genomes. The two confirmed fusion proteins originate from strains of *Microbacter sp.* and *Rhodococcus rhodochrous*. The latter of the two (CYP177A1) has been characterised

and shown to have activity against RDX (royal demolition explosive) allowing the host organism to mineralise the explosive while using it as sole nitrogen source[180]. The enzyme has been given the trivial name XplA after this activity and its redox partner, an enzyme resembling adrenodoxin reductase, is named XplB.

For the remainder of the potential P450-redox partner fusions in the table, no efforts towards characterisation are reported in the literature. For this reason they are not confirmed examples of novel fusion classes.

The fourth P450 fusion in this category consists of a P450 haem domain fused with a benzoate dioxygenase reductase (BDOR) system. This type of fusion is completely uncharacterised and therefore its nature is unconfirmed. However, BDOR belongs to the same superfamily of dioxygenase reductases as PDOR, and could potentially fulfil the same role as a redox partner to a P450 enzyme[181]. This suggests that the domain configuration could represent a further novel class of P450-redox partner fusion enzyme. An unusual feature of this domain architecture is the encoding of a DNA-binding domain between the reductase and P450 domains, though this does not preclude the system functioning as a self-sufficient P450 enzyme.

The fifth P450 fusion combines an N-terminal NAD(P)H-binding domain based on the Rossmann fold, with a C-terminal P450 haem domain. Potentially, parts of this fusion protein could abstract electrons from the bound NAD(P)H and mediate the reduction of the P450 haem as part of its catalytic cycle. If so, this would constitute a novel redox partner class.

4.1.1.2. P450-Known Enzyme Fusions

The majority of the results returned by the CDART tool fall into this category, they consist of P450 haem domains on the same protein as other enzyme domains with functions predicted from homology to other enzymes. For these P450 fusions, it appears likely that the enzyme domains present have a related function to the P450. Unlike the redox partner fusions, these would more likely exist to enable substrate channelling (or at least efficient transfer) between adjacent enzymes in a pathway. P450 fusions of this type are presented in Table 4.2.




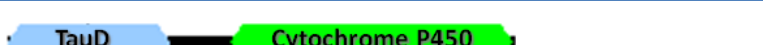



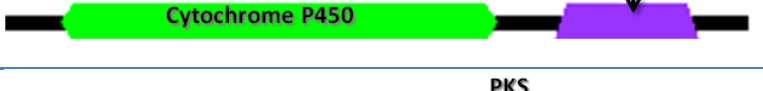

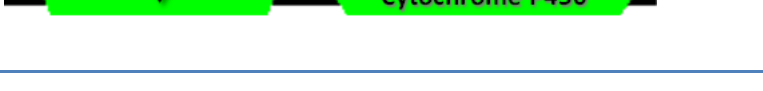
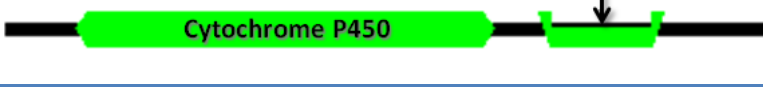
Fusion type	Layout	Number of sequences
Carboxyl esterase		2
Cinnamyl alcohol dehydrogenase		2
Possible α -dioxygenase		95
Taurine catabolism dioxygenase		2
3,4-dihydroxyphenylacetate 2,3-dioxygenase		2
Dyp peroxidase		7
Pectate lyase		2
Polyketide synthase		15
Polyketide synthase acyl-transferase domain		2
Ubiquinone/menaquinone C-methyltransferase		1
α/β hydrolase		2

Table 4.2. Selected results from a CDART search using the amino acid sequence of P450 BM3 showing domain layouts of potential P450 fusions with domains homologous to known enzymes. Graphical representations of the domain layouts are annotated from the CDART tool[178, 179].

An example of the potential for related functions is the P450-Cinnamyl alcohol dehydrogenase (CAD) fusion. CAD is an enzyme that converts cinnamyl alcohol to cinnamaldehyde producing one molecule of NADPH from NADP⁺ in the process. It is possible that the fused P450 further oxidises the product of this enzyme to produce a

carboxylic acid, a reaction shown to be performed by hepatic microsomal P450s[182]. It is further possible that the reducing equivalent produced by CAD is used by the P450, either directly or more likely through a separate redox partner.

An additional P450-known enzyme fusion that was not returned by the CDART search but has been previously mentioned in the literature[84] is the P450-acyl-CoA dehydrogenase (ACAD) fusion from *Pseudomonas fluorescens* (see Figure 4.2) . This fusion protein, given the systematic name CYP221A1, also has the potential to be a redox partner fusion, as the ACAD domain contains a FAD cofactor which becomes reduced during the reaction with acyl-CoA. This cofactor in the ACAD domain could possibly supply electrons to the P450 haem domain to which it is fused.



Figure 4.2. Domain schematic for the P450-acylCoA dehydrogenase (ACAD) fusion (CYP221A1) from *Pseudomonas fluorescens*.

In these as yet unproven models, an elegant system is suggested whereby one enzyme reaction in which electrons are incidentally abstracted is used to drive another reaction that requires electrons in a neighbouring fused enzyme domain *i.e.* a P450 reaction. The two reactions might even be subsequent steps in a metabolic pathway in which the external input of electrons is not required due to this ‘energy-saving’ measure.

4.1.1.3. P450-Other Domain Fusions

The remaining P450 fusions (presented in Table 4.3) do not appear to be combinations of P450s with redox partners or other enzymes, but rather with other recognised domain structures or short repeating sequences.

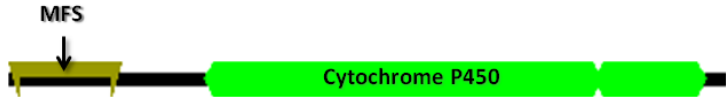

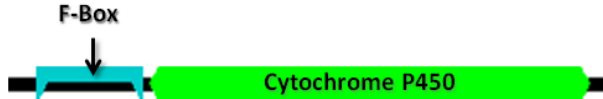



Fusion type	Layout	Number of sequences
Major facilitator superfamily		2
Fascin		2
F-Box type domain		2
Tetratricopeptide repeat etc.		2
Pentatricopeptide repeat		2
RING-type Zinc finger (RING finger)		3

Table 4.3. Selected results from a CDART search using the amino acid sequence of P450 BM3 showing domain layouts of potential P450 fusions with domains homologous to those involved in protein-protein interactions. Graphical representations of the domain layouts are annotated from the CDART tool[178, 179].

The major facilitator superfamily is a superfamily of transporter proteins (uniporters, symporters and antiporters) which facilitate the transport of a range of substrates across membranes[183]. It is possible in this fusion that either the P450 reaction product is immediately transported across a membrane (*e.g.* to prevent a build up), or that the substrate conveyed across the membrane must immediately undergo a reaction with the P450 (*e.g.* to reduce cytotoxicity).

Fascin is a domain involved in the bundling and cross-linking of actin filaments to produce cellular structures involved in the cellular matrix and architecture[184]. The genes encoding these putative fusions are found in strains of *Cyanothece sp.* a genus of cyanobacteria. It is uncertain why any such structures in these organisms would require the presence of a P450 haem domain.

The remaining putative fusion proteins contain motifs known to mediate protein-protein interactions. In these cases it may be that an enzyme complex is formed around the P450 domain facilitated by the motif in question.

The F-box motif is known to be involved in protein-protein interactions, and has been identified in many eukaryotic organisms where it is involved in the formation of an enzyme complex in the ubiquitination pathway[185]. Both of the F-box fusions are found in strains of the plant species *Arabidopsis thaliana*.

The next two examples in the table contain repeated peptide sequences, tetratricopeptides and pentatricopeptides. These 34 kDa (35 kDa) sequences are involved in recognition for protein-protein interactions in higher eukaryotes[186, 187].

The final putative fusion protein contains a zinc finger domain known as a RING (really interesting new gene) finger. This protein binds two atoms of zinc in a so-called 'cross-brace' pattern. Though the function of the RING domain is as yet unknown, many examples have been identified in organisms from plants to humans associated with diverse proteins[188].

4.1.2. CYP116B1

The first member of the CYP116 family to be identified and characterised was CYP116A1 from *Rhodococcus erythropolis*. This organism was shown to have a biosafening effect on crop plants against the application of the thiocarbamate herbicides EPTC (*S*-ethyl dipropylthiocarbamate) and vernolate (*S*-propyl dipropylthiocarbamate). An EPTC-induced gene cluster, was identified in the organism, which contains genes coding for a cytochrome P450 enzyme (*thcB*) and, downstream, a [2Fe-2S] ferredoxin named rhodocoxin (*thcC*) and a ferredoxin reductase (*thcD*)[119]. These genes together constitute the full P450/redox partner system. The expressed gene products of these three genes, ThcB (CYP116A1), ThcC and ThcD were shown to be the origin of the thiocarbamate degrading activity of the *R. erythropolis* organism. This enzyme system was additionally found to convey the ability to degrade another herbicide with a significantly different structure, atrazine[118].

Later, an enzyme was discovered in *Rhodococcus sp.* NCIMB 9784 which was found to represent a new class of self sufficient P450 enzyme; a P450-reductase fusion[189]. Named P450-RhF, the enzyme was found to contain a P450 haem domain of the CYP116 family fused to a reductase domain resembling phthalate dioxygenase reductase (PDOR). This reductase system consists of ferredoxin (Fd)-like and ferredoxin reductase (FdR)-like domains, the former containing a plant-type [2Fe-2S] cluster and the latter an FMN cofactor. From structural analysis of PDOR from *Pseudomonas cepacia*, the FdR-like domain has been shown to be composed of two subdomains, one binding the FMN cofactor and the other containing the binding site for a NAD(P)H coenzyme[132]. While the individual domains constitute the elements of a classic class I P450 redox partner system, the homology with PDOR and the fused domain arrangement made this enzyme the first example of a novel P450 redox partner system[190]. A study of P450-Rhf, later given the systematic classification CYP116B2, has confirmed the presence of FMN and [2Fe-2S] cofactors in the isolated Fd and FdR domains. The same study showed that the full length enzyme was monomeric and determined redox potentials for both the intact enzyme and isolated domains of CYP116B2[191]. Further work on CYP116B2 has demonstrated hydroxylation and *O*-dealkylation of several alkyl aryl ethers and highlighted a preference of the enzyme for shorter alkyl groups in these substrates[192].

Further members were classified as belonging to the A and B subfamilies of CYP116 from other *Rhodococci* with the ability to metabolise alkyl ethers[193].

The related enzyme CYP116B1 was identified in the genome of *Cupriavidus metallidurans* (named *Ralstonia metallidurans* at the time). The *CYP116B1* gene codes for the same domain structure as P450-Rhf (*i.e.* P450 haem and phthalate dioxygenase reductase-like domains). Interestingly, this organism contains genes coding for fusion P450 enzymes of both the P450-Rhf type (CYP116B1) and the BM3 type (CYP102E1).

CYP116B1 does not have a known physiological function. However, the relationship with CYP116A1 and CYP116B2 predicts an activity against thiocarbamates and potentially atrazine[118, 119]. Indeed, hydroxylation of two thiocarbamate herbicides, EPTC and vernolate, has been shown for CYP116B1[135].

In this thesis the work on CYP116B1 was on two fronts: The characterisation of CYP116B1 and isolated domains thereof as proteins (*i.e.* expression, isolation and biophysical properties), covered in this chapter, and of its enzymatic properties which are covered in chapter 5. The focus of the protein work was thermodynamic, spectroscopic and structural characterisation of the full length protein and its isolated haem and reductase domains. This began with establishing a protocol for efficient expression and purification of these protein components.

4.2. Cloning

4.2.1. FLCYP116B1

A protocol for the successful expression and purification of CYP116B1 had previously been devised by Warman *et al.* [135]. For this protocol the gene coding for CYP116B1 had been cloned into the pET11a expression vector (New England Biolabs, NEB). This vector was transformed into the Origami (DE3) strain of *E. coli* (Novagen) which was then grown and induced overnight at 18 °C. This produced expression of soluble, intact CYP116B1 at a level, shown by SDS-PAGE, to be less than that of some constitutively expressed proteins in the *E. coli* strain. The crude CYP116B1 protein then required purification with three chromatography columns. Attempts to reproduce this protocol showed that it was not ideal due to a protracted purification during which proteolysis occurred and could not be prevented.

To try to improve recovery of CYP116B1 protein, an N-terminal 6xHis-tag was added by cloning into an alternative expression vector, pET-15b (NEB), which codes for six histidine residues and a short linker region before the gene for expression. This tag system could enable more efficient purification of CYP116B1 using affinity for nickel- or cobalt-containing resins. The vector also carries a gene for ampicillin resistance to allow for selection of cells carrying the plasmid.

The starting point for cloning into pET-15b was an available construct of the CYP116B1 gene, with appropriate initiation and terminations sequences, in the pGEM-T vector (Novagen). This construct was produced by Warman *et al.* for the purposes of creating a pET-11a clone containing the *CYP116B1* gene and also incorporates *AseI* and *BlnI* restriction enzyme sites at either side of the gene to

facilitate subcloning (Figure 4.3)[135]. The AseI site includes the start codon for the gene and the BlnI site is some distance after the stop codon, as the insert includes some non-coding DNA from the *C. metallidurans* genome.

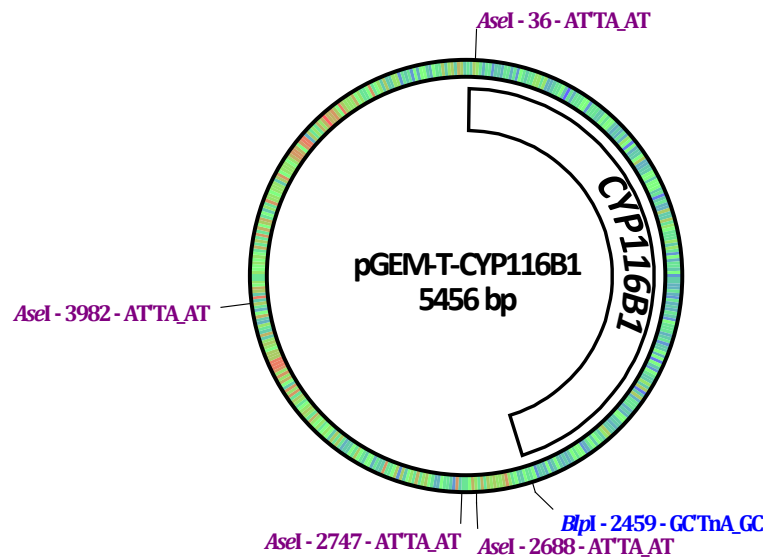


Figure 4.3. Plasmid map of a pGEM-T construct containing a PCR fragment of the gene encoding CYP116B1 (Rmet4932) from the genome of *Cupriavidus metallidurans*. This construct was created previously by Warman *et al.*[135] who engineered AseI and BlnI restriction sites at the start and end of the gene, respectively, for the purposes of subcloning.

The pET-15b vector contains a T7 promoter and a sequence coding for a 6xHis tag immediately upstream of the multiple cloning region. In this region are NdeI and BlnI restriction enzyme sites required for insertion of the *CYP116B1* gene. Further upstream is the *lacI* gene, the repressor from the lac operon, which allows expression of the cloned gene to be controlled using lactose-type inducers such as allolactose and IPTG. The plasmid also carries the gene coding for β -lactamase, conferring resistance to ampicillin and related antibiotics. This allows for antibiotic selection to be used to encourage retention of the vector by the host strain. (see Figure 4.4)

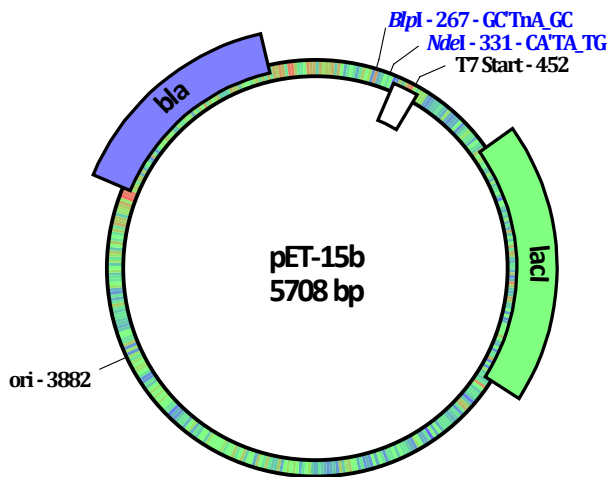


Figure 4.4. Plasmid map of the vector pET-15b supplied by NEB. Indicated, within the cloning region, are the BlnI and NdeI restriction enzyme sites used in the cloning described in the text. In blue is the gene coding for β -lactamase (labelled *bla*) and in green is the *lacI* repressor gene. The white label indicates the T7 promoter region and the sequence coding for a 6xHis tag. Additionally, the position of the start of transcription is indicated.

The pGEM-T construct was prepared on a suitable scale from *E. coli* transformants using a midi prep kit (QIAGEN) as described (section 2.4.5) and the resulting plasmid stock was double digested with AseI and BlnI restriction enzymes to excise the gene coding for CYP116B1 from the vector. The excised gene insert was separated from the other fragments of the pGEM-T vector by running the digested sample on a 0.8 % agarose gel containing ethidium bromide.

A sample of the pET-15b vector containing an existing gene insert was cut with NdeI and BlnI restriction enzymes. The linearised vector was separated from the unwanted excised insert by running the digested sample on a 0.8% agarose gel containing ethidium bromide.

The gels were visualised using a UV transilluminator (Figure 4.5) and the bands corresponding to the *CYP116B1* gene insert and linearised pET-15b vector were carefully cut from the gel and the DNA within them was purified using a DNA gel extraction kit (QIAGEN).

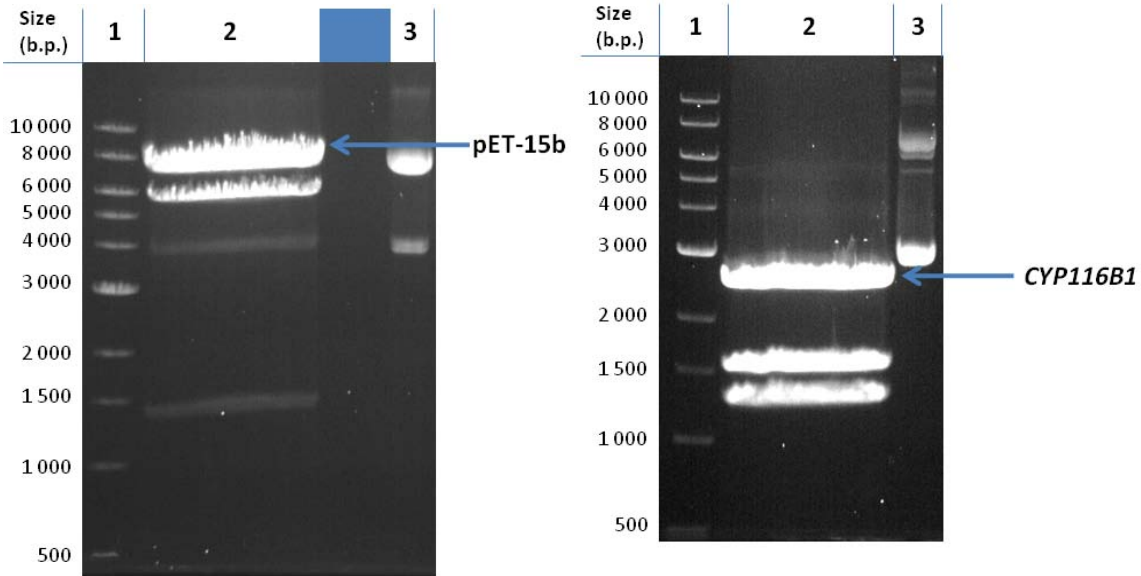


Figure 4.5. Images of 0.8 % agarose gels showing pET-15b vector (with a gene insert) double digested with NdeI and BlnI restriction enzymes (left) and pGEM-T-CYP116B1 double digested with AseI and BlnI enzymes (right). In both cases the lanes show NEB 1kb DNA ladder (1), the digested DNA (2) and an undigested sample for comparison (3). The vector digest appears to be successful as the gel shows two bands representing the original insert and the linearised vector. However, these DNA fragments do not appear to have travelled an appropriate distance on the gel compared to the bands of the DNA ladder. The insert digest is also successful; the gel shows the *CYP116B1* insert (2360 b.p.) at an appropriate position along with two shorter fragments of the pGEM-T vector, which is cut five times in total to produce three long fragments and two short fragments.

NdeI and AseI restriction enzymes are not isoschizomeric, but produce cohesive ends which can be ligated (Figure 4.6).

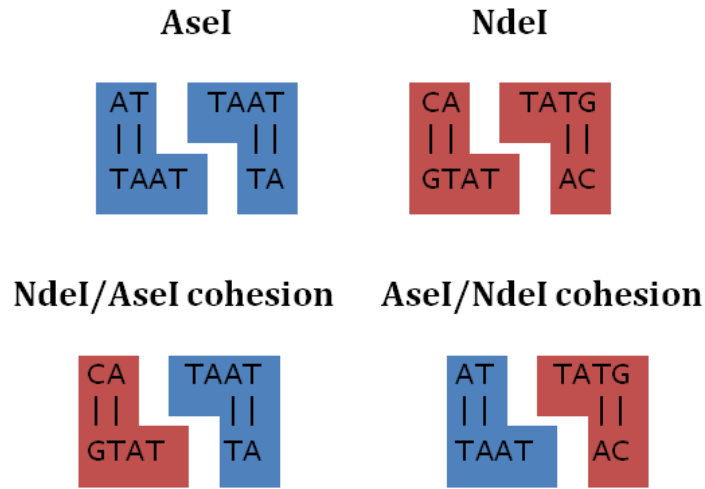


Figure 4.6. Diagram showing the recognition sites and cutting patterns for AseI and NdeI restriction enzymes. The bottom diagram shows how AseI and NdeI sticky ends can be cohesive, allowing them to be ligated. The ATG from the NdeI site which forms the start codon of the CYP116B1 gene is retained in this ligation.

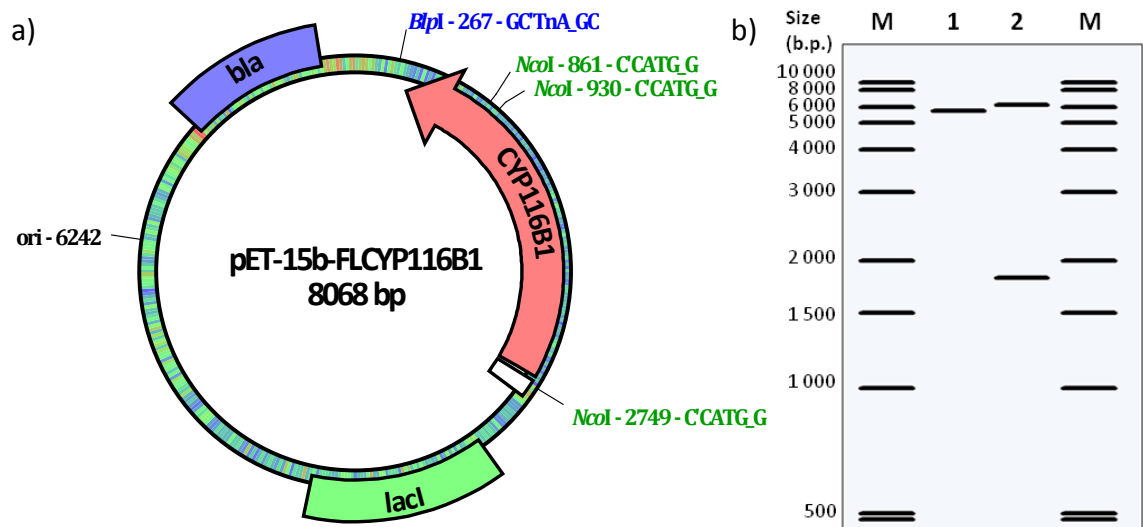


Figure 4.7. a) Plasmid map of the pET-15b construct containing the insert of the CYP116B1 gene from *C. metallidurans*. The gene insert is shown in red and all other genes are labelled as in the pET-15b plasmid map (Figure 4.4). Also shown are the NcoI restriction sites used to confirm the integrity of the construct after cloning and the reconstituted BplI site after the gene. b) Simulated agarose gel showing the predicted cutting pattern for pET-15b vector (1) and pET-15-FLCYP116B1 clone (2).

Ligation reactions were set up as described previously (section 2.4.2) with varying ratios of vector to insert DNA (1.5:1, 1:1, 1:3 and 1:5) in the presence of T4 DNA ligase enzyme. This was achieved using a constant amount of vector DNA and varying the amount of insert DNA. In addition, a control reaction was set up for each ratio of vector to insert, wherein the DNA ligase enzyme was excluded.

Each reaction was used to transform NovaBlue cells which, when plated out on to ampicillin-containing growth medium, produced varying numbers of colonies. For each reaction two colonies were selected (designated A and B), grown out and the plasmid DNA within was purified.

Each of the potential new pET-15b constructs was analysed by restriction digest with the NcoI restriction enzyme. The pET-15b vector sequence contains one NcoI site and the gene insert contains two, although they are close together in the sequence (see Figure 4.4 Figure 4.7a). The pET-15b-FLCYP116B1 construct (containing the intact *CYP116B1* gene) digested with NcoI would be cut into three fragments of 69, 1819 and 6180 base pairs. A 0.8 % agarose gel of the digested plasmid would show two bands, as the shortest of these fragments is likely to migrate too quickly and run off the gel. By comparison, the pET-15b vector without the insert would be linearised and appear as one band (see Figure 4.7b). The NcoI digested DNA (along with an undigested sample for comparison) was run on a 0.8 % agarose gel (Figure 4.8). All samples for the reactions using the two highest ratios of vector to insert (1.5:1 and 1:1) showed the predicted pattern of bands as described above. DNA from reactions with lower ratios showed unexpected results indicating that the vector DNA present was not the intended construct.

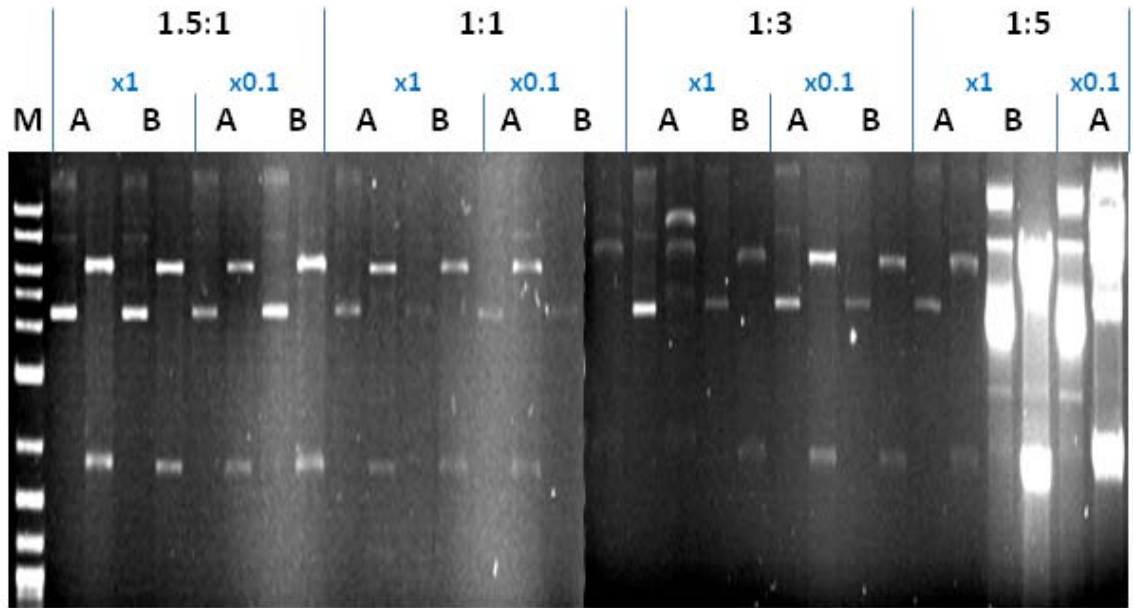


Figure 4.8. Agarose gel (0.8%) showing miniprep DNA (*NcoI* undigested and digested) from ligation reactions aimed at creating a pET-15b-FLCYP116B1 plasmid construct. The first lane shows a DNA ladder (1 kb ladder, NEB). Proceeding lanes show miniprep DNA (1 μ L) from NovaBlue cells transformed with ligation reactions which are labelled by the ratios of vector to insert DNA used (1.5:1, 1:1, 1:3 and 1:5). Each ligation reaction was used to transform the cells from both undiluted (x1) and diluted one in ten (x0.1) stocks. Each of these transformations was plated out onto antibiotic-containing medium and two of the resultant colonies selected (denoted A and B). For each DNA sample shown, the left hand lane shows the undigested DNA and the right hand lane shows the *NcoI* digested sample. All lanes for the two highest vector to insert ratios (1.5:1 and 1:1) seem to show the predicted pattern of bands for their *NcoI* digests (Figure 4.7b), suggesting they contain the desired pET-15b-FLCYP116B1 construct.

The integrity of the plasmid constructs in the cells was also confirmed by DNA sequencing which revealed an unforeseen sequence error. The ligation of the *AseI* cut vector and the *NdeI* cut gene fragment had created an insertion of a single adenosine nucleotide before the first codon in the *CYP116B1* gene (see Figure 4.9)

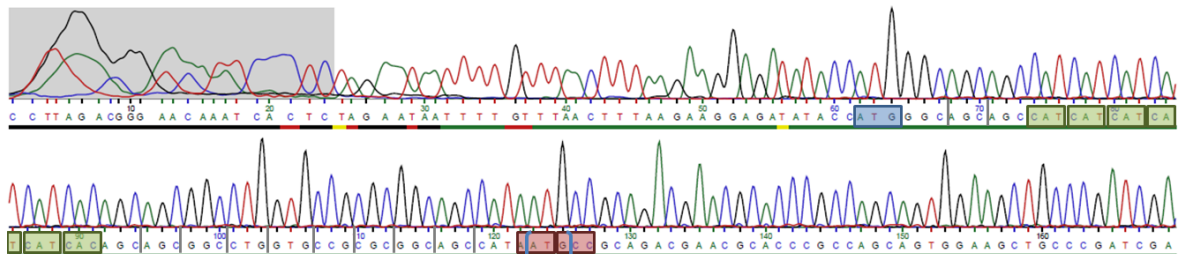


Figure 4.9. DNA Sequencing chromatogram showing the pET-15b start codon (blue), the six codons coding for the 6xHis tag (green) and the area around the intended start codon (blue brackets), where an extra adenosine nucleotide from the *AseI*/*NdeI* site cohesion shifts the sequence out of frame. The first two codons of the nonsense protein are also indicated (red).

This error would have also occurred in the original pET-11a construct, however it would not necessarily have prevented proper expression as there is some tolerance for variation in the distance between upstream promoting sequences and the start codon[2]. However, it may explain, to some degree, the limited levels of expression that were achieved with this construct if this distance was not optimal.

The extra nucleotide presented a more serious problem in pET-15b, where an earlier initiation site encompassing the 6xHis tag is used. In this case the extra nucleotide falls at the end of the linker between the region encoding the tag and that encoding the CYP116B1 protein, and is translated faithfully. This causes a frame shift which would lead to the translation of a nonsense protein after the 6xHis tag.

The extra nucleotide at the 61st position (after the pET-15b start codon) was deleted by site directed mutagenesis as described (see section 2.4.1). The primers used were FL61delFwd (5'-GCCGCGCGGCAGCCAT*ATGCCGAG-3') and FL61delRev (5'-CTGCGGCAT*ATGGCTGCCGCGGC-3') with * indicating the site of the deletion.

4.2.2. HDCYP116B1

A separate construct for the expression of the isolated haem domain of CYP116B1 (HDCYP116B1) was created from the full length construct by inserting two stop codons into the gene at the linker region between the haem and reductase domains. This was achieved by site-directed mutagenesis of the pET-15b plasmid containing the gene for CYP116B1, using the method previously developed by Warman *et al.* to create a pET-11a-HDCYP116B1 construct from a pET-11a-FLCYP116B1 construct[135]. This method involved the creation of two in-frame stop codons in the vector DNA after the region coding for the haem domain. This truncates translation before the reductase domain, meaning only the haem domain is expressed. The primers used were the same as in the work by Warman *et al.* - HDmutFwd (5'-GACGGTGCTGGCGCCA[TGATAG]GCGGTCCGCATCGG-3') and HDmutRev (5'-CCGATGCGGACCGC[CTATCA]TGGCGCCAGCACCGTC-3') with nucleotides in square brackets representing the two stop codons to be inserted. These primers bind to a part of the *CYP116B1* gene identified as coding for the linker region between the haem and reductase domains (1335 – 1371 b.p.) where the stop codons are to be

inserted. The successful creation of stop codons was confirmed by internal sequencing of the *CYP116B1* gene in the pHDCYP116B1-15b candidate constructs. The sequencing used an internal sequencing primer EndHDseqF (5'-CGATCTCGAAGGCGCTGGTATCG-3') which binds the *CYP116B1* gene at the end of the region coding for the haem domain (1566 – 1588 b.p.) (Figure 4.10).

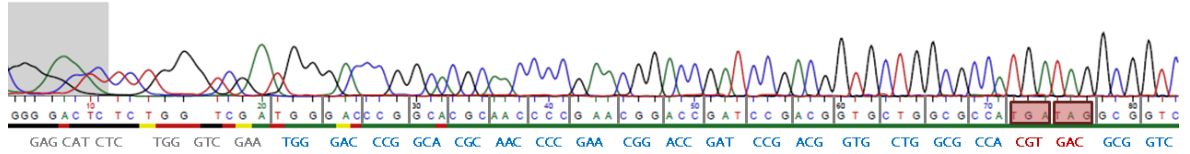


Figure 4.10. Beginning of sequencing chromatogram for the pHDCYP116B1-15b candidate, showing the end of the region coding for the haem domain and the successfully introduced stop codons (red boxes). Shown below the chromatogram is the *CYP116B1* gene sequence divided into codons in the proper reading frame. This highlights the integrity of the sequence (where the sequencing data are reliable, in blue) and the two mutated codons (in red).

Based on the sequencing data, the candidate construct was confirmed to be the desired pHDCYP116B1-15b clone.

4.2.3. RDCYP116B1

An existing pET-11a construct containing the portion of the *CYP116B1* gene coding for the reductase domain (pRDCYP116B1-11a) had been created by Warman *et al.*[135]. Sequencing to confirm its integrity revealed that this construct contained a point deletion of a single guanine nucleotide in its DNA sequence.

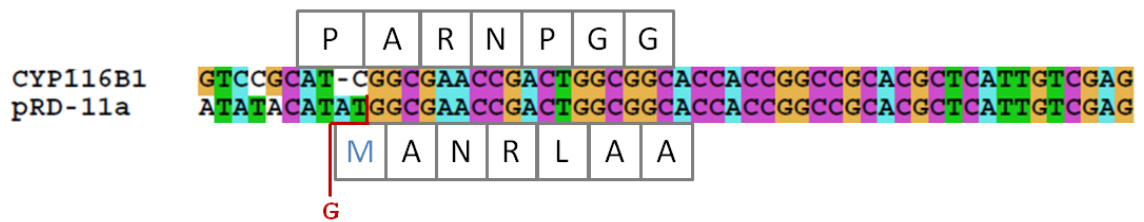


Figure 4.11. Section of an alignment of the *CYP116B1* gene sequence (1 360 – 1413 b.p.) and the resulting sequence from DNA sequencing of the pRDCYP116B1-11a construct. This shows the upstream sequence that was introduced in the construct to allow expression of the isolated portion of the gene coding for RDCYP116B1 (TATACAT). This is followed by the start codon (ATG) where the position of the deleted guanine nucleotide is indicated (red). By chance, the first nucleotide of the next codon is also guanine which restores the start codon but also alters the reading frame of translation. This is demonstrated by showing the start of translation for two sequences, which reveals the expression of a nonsense protein for the pRDCYP116B1-11a sequence.

Many samples of the construct were sequenced and all were found to contain the deletion. This suggests that the mutation probably occurred due to either a reading error or a mutated primer in the original PCR reactions carried out during the creation of the construct. The deletion would shift the gene sequence out of its natural reading frame leading to the expression of a 125 residue nonsense protein instead of RDCYP116B1 (325 residues) (see Figure 4.11 & Figure 4.12).

<p>MANRLAAPPAARSLSSAS RRPRRVCRGSAWFRPMVA HYRAGRARRISTSNAATP ASRASIRCAATPPIPAPS RSRCCASP KAVVDRRGFM PACAQATSSRCAARAIT S GSTRPAVARSSSPVALA_</p>	<p>MGEPTGGTTGRTLIVERVETAA QGVSRIRLVSPDGRALPRWSPG SHIDIECGHTGISRQYSLCGDP ADTSAFEIAVLREPE SRGGS AW IHASLRAGDKLKVRGPRNHFR L DETCRRAIFIAGGIGVTPVSAM ARRAKELGVDYTFHYCGRSRAS MAMIDELRALHGDRVRIHA ADE GQRADLAQVLGAPDANTQIYAC GPARMIEALEALCATWPEDSLR VEHFSSKLGTLDP SREQPFAVE LKDSGLTLEVPPDQTL LALRA ANIDVQSDCEEGLCGSCEVRVL AGEIDHRDVVLTRGERDANNRM MACCSRAAKGGKIVLGL_</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figure 4.12. Side by side comparison of the sequences of the protein that would be expressed from the pRDCYP116B1-11a construct with the out-of-frame error described in the text (red) and that of the intended expression *i.e.* the correct RDCYP116B1 protein (blue). The sequence of the out-of-frame protein bears very little resemblance to that of RDCYP116B1 and terminates after 125 residues (*c.f.* 325 residues for RDCYP116B1) due to an in-frame stop codon (represented by an underscore).

The deleted nucleotide was reinserted by site-directed mutagenesis as described (section 2.4.1). The primers used were RDinsertG2Fwd (5'-GAAGGAGATATACATATG[G]GCGAACCGACTGGC-3') and RDinsertG2Rev (5'-CTTCCTCTATATGTATAC[C]CGCTTGGCTGACCG-3'). The inserted nucleotide is indicated in the sequences by square brackets. These primers bind to the pRDCYP116B1-11a vector DNA containing the out-of-frame mutation, spanning the site to be mutated and the start codon (ATG) (-15 - +18) on the sense and antisense strands.

The correction of the out-of-frame mutation was confirmed by sequencing using a standard T7 forward primer (provided by Eurofins MWG Operon). The chromatogram clearly shows the restoration of the proper sequence with the insertion of the missing guanine nucleotide (Figure 4.13).

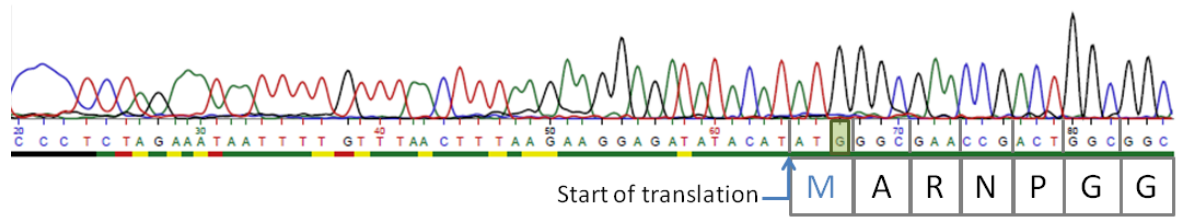


Figure 4.13. Section of the sequencing chromatogram showing the T7 forward sequence of the pRDCYP116B1-11a construct with the out-of-frame error corrected by the addition (as described in the text) of a single guanine nucleotide in the 3rd position after the start of translation (green box). The now correct translation of the start of the RDCYP116B1 protein is shown below the chromatogram (c.f. sequences in Figure 4.12 & Figure 4.11).

4.3. Expression and Purification

4.3.1. FLCYP116B1 and HDCYP116B1

Expression trials were conducted as described in section 2.4.6. The aim of the trials was to determine the cell strain providing the best yield of CYP116B1 and the haem domain of the enzyme (HDCYP116B1). SDS PAGE gels of the expression trials show a band at ~89 kDa, the correct mass for CYP116B1, for a number of the strains tested (Figure 4.14). Three strains, BL21 (DE3), HMS174 (DE3) and Rosetta 2 (DE3), all seemed to provide excellent levels of expression (with IPTG induction at 500 μ M). In these strains the start of expression of FLCYP116B1 can be seen in the samples taken after two hours, with peak expression seen in the samples taken after twenty hours. No further increase in expression is shown for the samples taken after twenty four hours. For both HMS174 (DE3) and BL21 (DE3) strains, expression is particularly good, with an intense band appearing on the gel at a weight corresponding to FLCYP116B1, compared to the other bands representing constitutively expressed *E. coli* proteins.

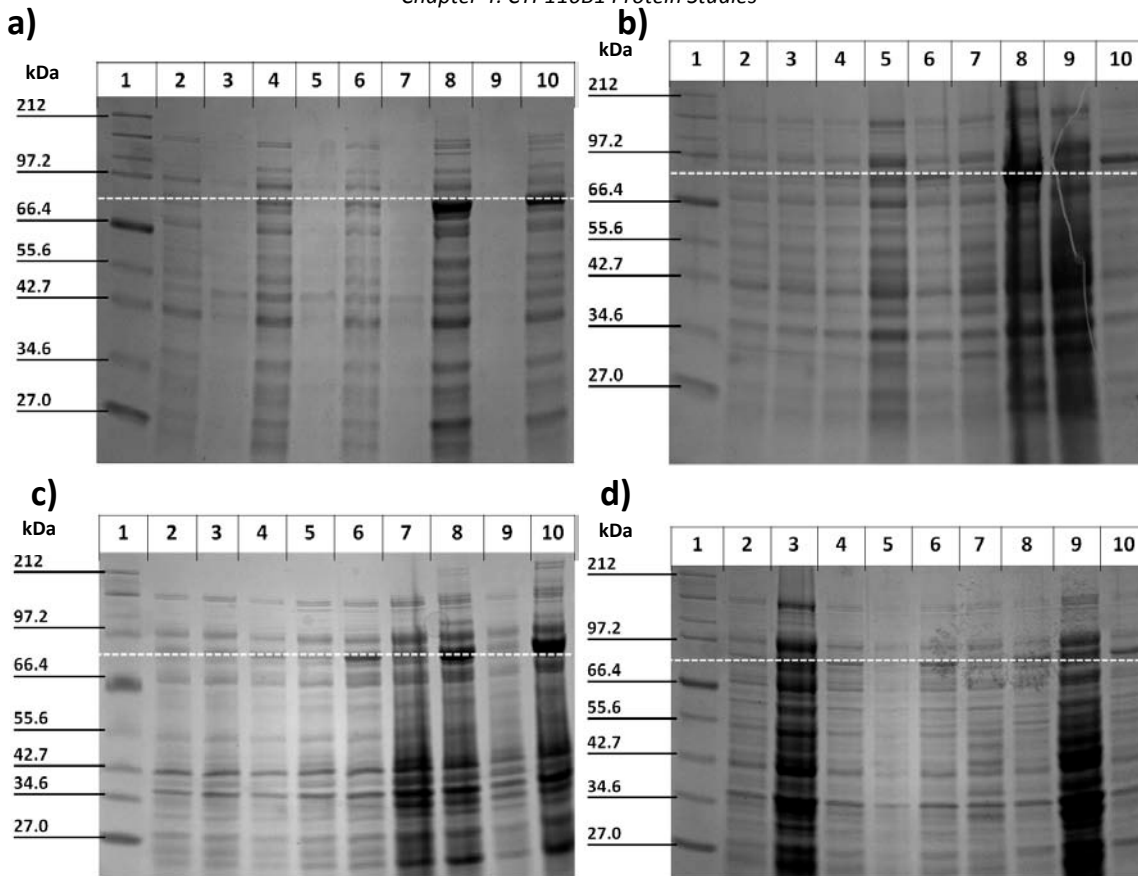


Figure 4.14. Images of SDS PAGE gels showing samples of expression trial cultures at time intervals with and without IPTG induction. Each gel is for a different expression strain a) HMS174 (DE3) b) BL21 (DE3) c) Rosetta2 (DE3) d) Origami (DE3) (all Novagen). In each gel the lanes are as follows: 1) Broad range protein marker (NEB) 2) Culture sampled at the point of induction. (t_0). 3) Uninduced culture sampled 1 hour after t_0 . 4) As lane 3 but from induced culture. 5) Uninduced culture sampled 2 hours after t_0 . 6) As lane 5 but from induced culture. 7) Uninduced culture sampled 20 hours after t_0 . 8) As lane 7 but from induced culture. 9) Uninduced culture sampled 24 hours after t_0 . 10) As lane 9 but from induced culture. A white dashed line indicates the position of bands due to FLCYP116B1. Each culture sample loaded onto the gel was normalised based on the OD_{600} of the culture at the time of sampling as described in the pET system manual[194] and 10 μ L were loaded in each case.

Following expression trials, a 50 mL culture of each of the best performing strains was grown and induced with IPTG (as during the trials) and then the cells were pelleted in 50 mL tubes by centrifugation at 4500 x g in a benchtop centrifuge. Pellets were resuspended in 10 mL phosphate buffer, protease inhibitors PMSF and benzamidine were added and the tubes placed in an ice bath. Lysosyme and DNaseI were added and the cells were lysed was by sonication in six cycles of 30 seconds on, 60 seconds off at 40% power using a Vibracell sonicator with a CV-33 probe (Sonics & materials inc. Newtown, CT, USA.). Cell debris was removed by centrifugation at 40 000 x g and the supernatants were retained. UV/Visible spectrophotometric analysis of the supernatants revealed the highest concentration of CYP116B1 came from the HMS174 (DE3) culture.

Purification by 6xHis tag affinity was tested by loading the supernatant onto a small column (0.5 mL bed volume) of nickel-NTA resin in a syringe barrel. These small columns were washed with nickel purification buffer (section 2.2.2.3) and eluted by increasing concentrations of imidazole (30 – 250 mM). Fractions were analysed on a 12 % SDS-PAGE gel and the optimum elution of CYP116B1 was found to occur at imidazole concentrations between 50 and 100 mM. The purification trial also confirmed that the best yield was from the HMS174 (DE3) culture (equivalent to ~40 mg of FLCYP116B1 per litre of culture).

The expression and purification was scaled up with typical cultures of 14 L grown at a time. Typical yield after nickel purification was ~600 mg of FLCYP116B1 protein for one preparation. An SDS-PAGE gel of the purification process (Figure 4.15) shows that, after the initial flow through, relatively pure FLCYP116B1 is eluted at a range of imidazole concentrations. These fractions were pooled and the protein further purified by size exclusion chromatography. The full protocol can be found in section 2.4.8.

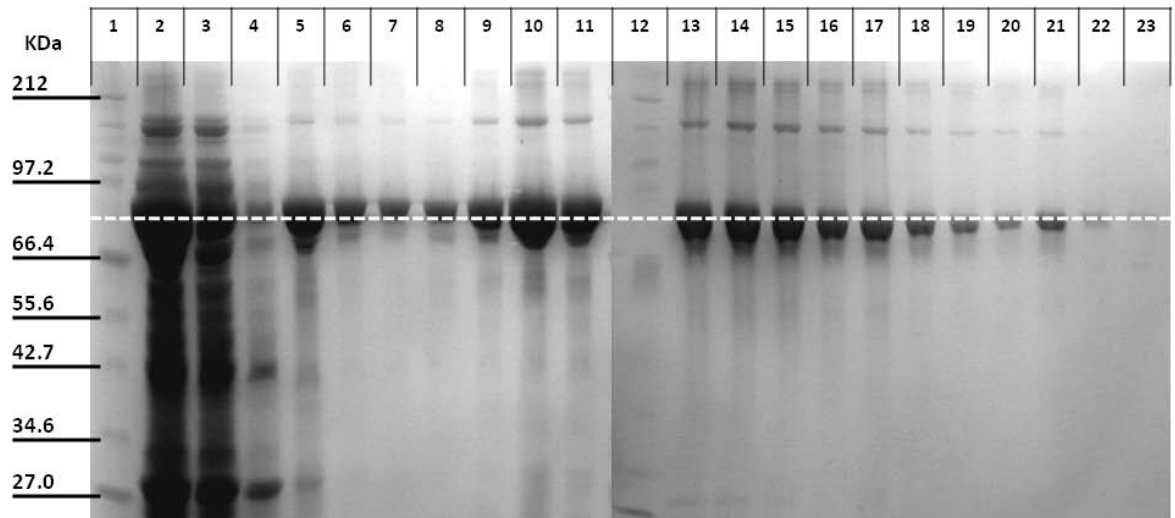


Figure 4.15 Image of an SDS-PAGE gel showing the purification of FLCYP116B1 using a 6xHis tag affinity column. Lanes are as follows: 1) Broad range protein marker (NEB). 2) Cell lysate containing FLCYP116B1. 3) flow-through from loading the column. 4) flow-through from washing the column with nickel buffer. 5) flow-through from washing the column with 20 mM imidazole in nickel buffer. 6-9) fractions from elution with 50 mM imidazole. 10-17) as 6-9 but with 100 mM imidazole 18-23) as 6-9 but with 250 mM imidazole. The dashed white line represents the molecular weight of FLCYP116B1.

HDCYP116B1 could be expressed and purified in the same way as FLCYP116B1 by transforming HMS174 (DE3) cells with the pHDCYP116B1-15b construct described

above (section 4.2.2). HDCYP116B1 was purified with comparable yields to FLCYP116B1.

It became apparent through UV/Visible spectroscopic studies that the haem of HD and FLCYP116B1 purified in this manner retained some imidazole as a bound ligand. This was true even after buffer exchange by size exclusion chromatography during purification and extensive dialysis. In order to maintain the levels of expression provided by the 6xHis tag purification scheme, while eliminating imidazole, an alternative method of elution from the nickel-NTA resin was devised using 100 mM EDTA to chelate nickel from the column and disrupt the interaction with the 6xHis tag. The nickel and EDTA were separated from the protein using centrifugal concentrators with a molecular weight cut off of 20 kDa. The method successfully produced a sample of HDCYP116B1 free from imidazole ligation, as evidenced by the UV/Visible spectrum and EPR (see section 4.5.2).

4.3.2. RDCYP116B1

The reductase domain of CYP116B1 (RDCYP116B1) was expressed from the HMS174 (DE3) strain of *E. coli* transformed with the pET-11a construct described above (section 4.2.3). The RDCYP116B1 protein was purified using a succession of three columns as described (section 2.4.9). Though a number of preparations of RDCYP116B1 protein were successful, the reproducibility of the purification was variable. A spectrum of the purified RDCYP116B1 is shown in Figure 1.16. This spectrum is quite similar to that recorded for PDOR from *Pseudomonas cepacia*[132].

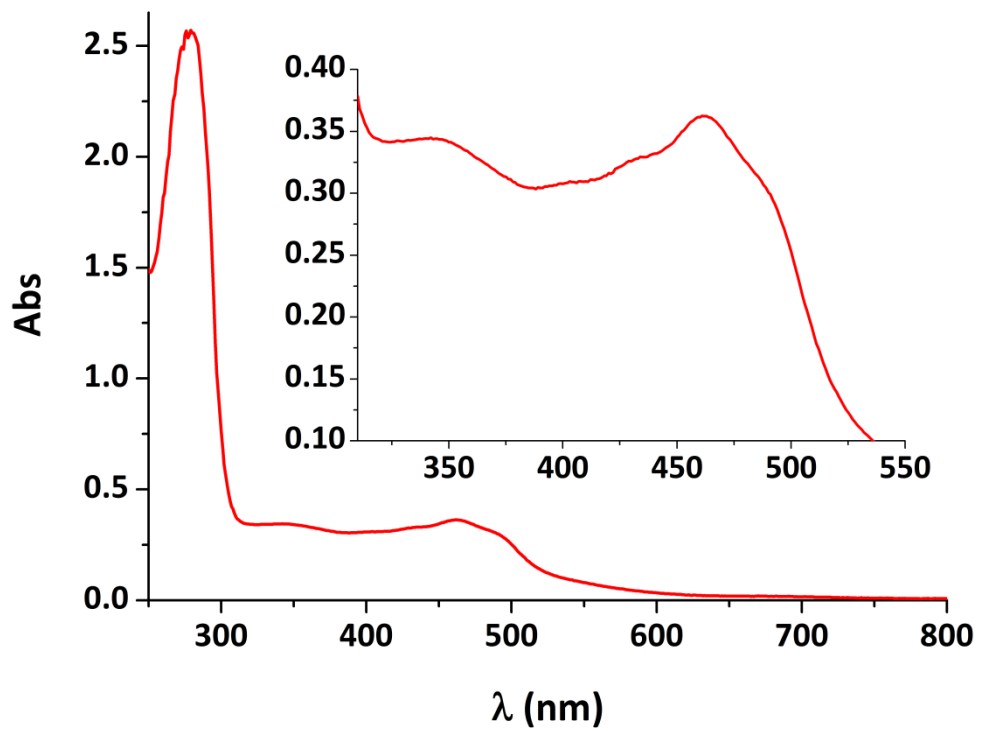


Figure 4.16 The UV/visible spectrum of purified RDCYP16B1 protein with an expansion of the region 310-550 nm as an inset. The spectrum closely resembles that of intact phthalate dioxygenase reductase[132]. An absorbance maximum for the FMN cofactor is observed at 463 nm with shoulders at 433 nm and 490 nm. As signal corresponding to the [2Fe-2S] cofactor is seen at 346 nm.

4.4. Binding of Small Ligands

The purified HDCYP116B1 and FLCYP116B1 proteins were characterised by the binding of small molecule ligands. The reduced, CO-bound UV/Vis absorption spectrum of FLCYP116B1 (Figure 4.17) shows a mixed population of P450 and P420 species. The reduced CO-bound spectrum of HDCYP116B1 was found to be virtually identical in this respect.

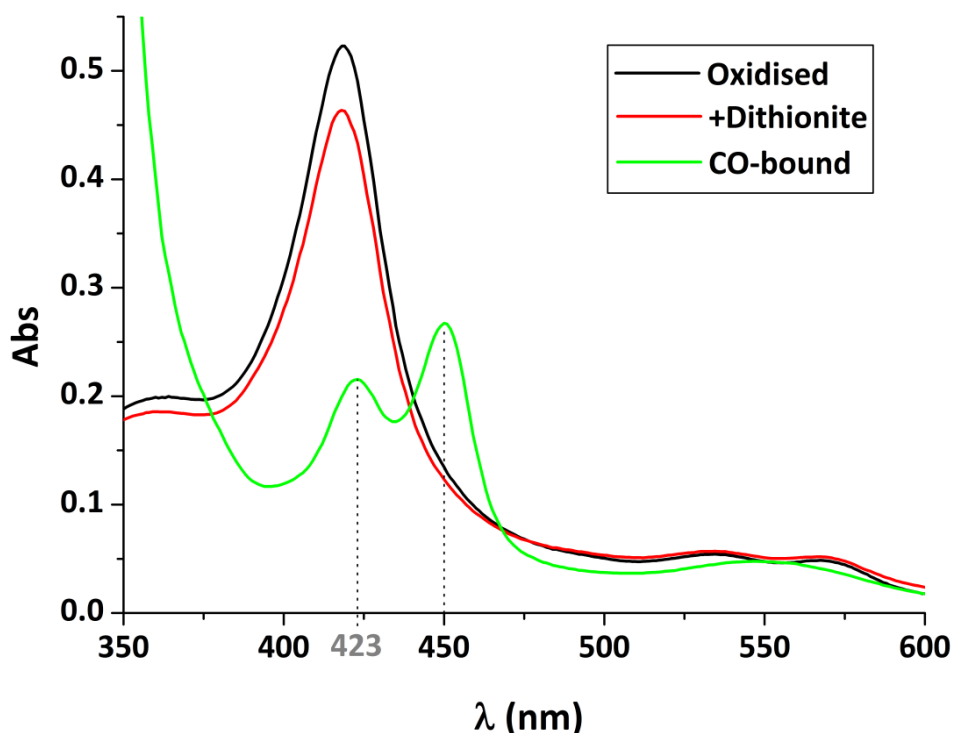


Figure 4.17. UV/Visible spectra of FLCYP116B1 ($\sim 5 \mu\text{M}$) in the oxidised state as purified (black), upon addition of solid sodium dithionite to 1 mg mL^{-1} (red), and reduced and in complex with CO (green). CO ligation of FLCYP116B1 haem results in a mixed population of P450 (peak at 450 nm) and P420 (peak at 423 nm) haem.

Compared to the parallel CO-binding spectra for HDCYP102A3, there appears to be a greater effect upon FLCYP116B1 from addition of the reducing agent sodium dithionite in terms of the effect on the Soret peak. This may suggest that a greater proportion of FLCYP116B1, than HDCYP102A3, is reduced by this step (*i.e.* the redox active cofactors in the reductase domain). In the reduced and CO-bound spectrum, two peaks are visible at 423 nm and 450 nm. These peaks are most likely the two CO-bound species known as P450 and P420, thiolate ligated carbonmonoxyhaem by the former and thiol-ligated carbonmonoxyhaem in the case of the latter. This result suggests that the haem-ligating cysteine residue of CYP116B1 (Cys-384) is more susceptible to protonation in the reduced state than that of HDCYP102A3 (Cys-402),

at least with respect to reduction of the haem iron by sodium dithionite. The ratio of P450 to P420 is 1.2 : 1 by absorbance values, although this is only a rough measure as it doesn't take into account any differences in molar absorbance coefficient between the two species. The P420 coefficient is typically larger than that for the P450 form, suggesting that the balance is in favour of the P450 form. The merging of the α/β bands at 534 nm and 567 nm to a single, broad spectral feature at 550 nm is characteristic of reduced P450 haem.

The binding of cyanide to the haem of HDCYP116B1 was examined by stepwise titration of a 1 M stock of NaCN into a sample of HDCYP116B1. After each addition the UV/Vis spectrum was recorded. The spectrum of HDCYP116B1 before addition of NaCN was subtracted from each subsequent ligand-bound spectrum to produce a series of difference spectra (Figure 4.18).

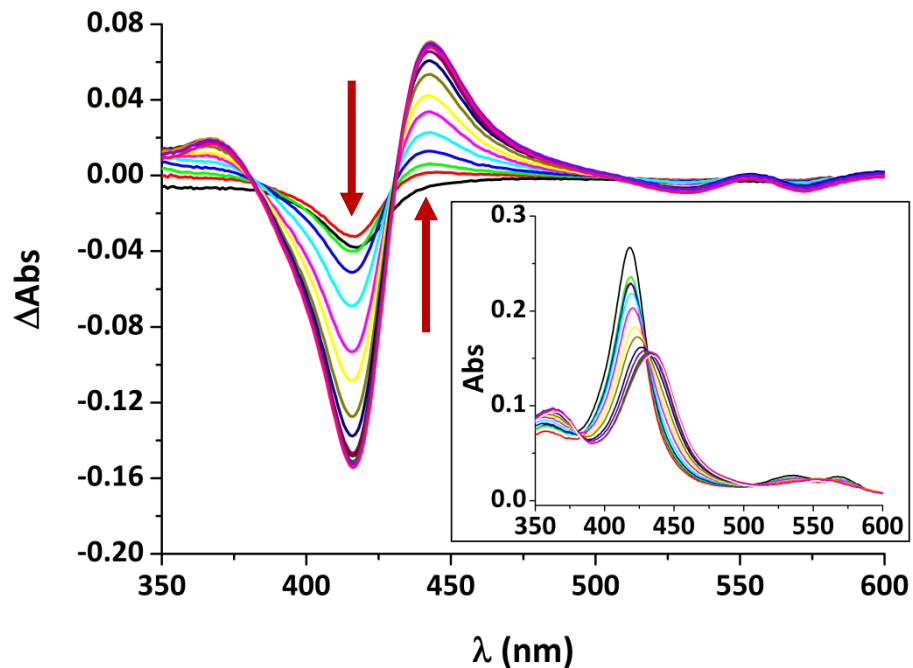


Figure 4.18. Difference spectra (with original spectra as an inset) showing the binding titration of cyanide, in the form of NaCN (200 – 2000 μM), against HDCYP116B1 ($\sim 3 \mu\text{M}$). A peak can be seen forming at 443 nm while a trough forms at 416 nm. Red arrows indicate the direction of absorbance changes upon cyanide addition. The Soret peak shift in the absolute spectrum is from 419 nm to 437 nm upon cyanide binding.

From the difference spectra, the wavelengths of maximal change were identified with a peak forming at 443 nm and a trough developing at 416 nm. The ΔAbs values recorded at 416 nm were subtracted from those at 443 nm and the resulting $\Delta\Delta\text{Abs}$ values were plotted against the concentration of NaCN to produce a binding curve

(Figure 4.19). The binding curve appeared sigmoidal and could be fitted using the Hill equation. The fit produced a value of the apparent dissociation constant (K_H) of 0.54 ± 0.02 mM and a Hill number of 3.00 ± 0.30 .

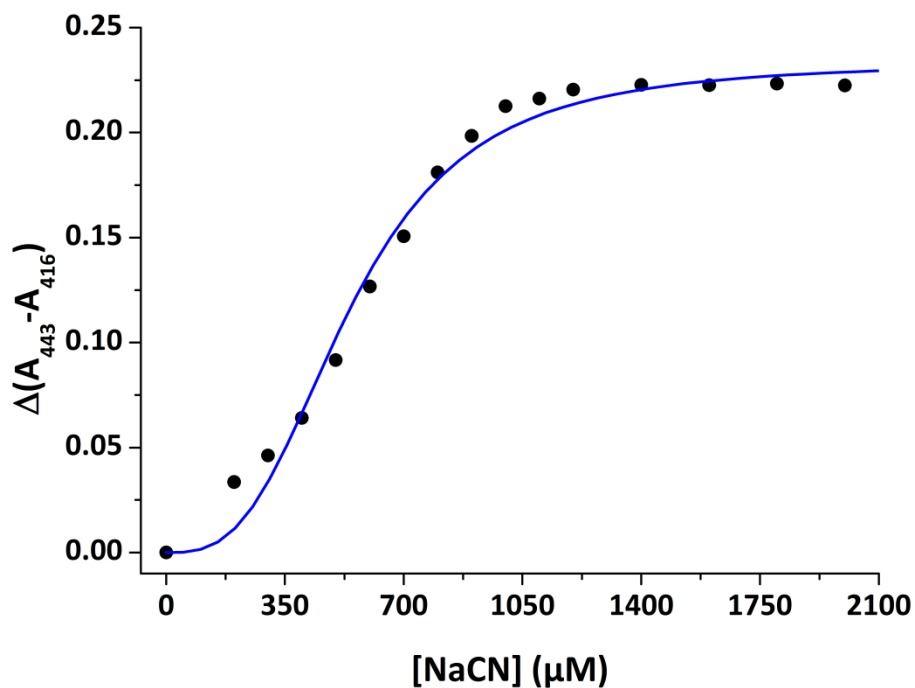


Figure 4.19. Binding curve for the titration of NaCN against HDCYP116B1 ($\sim 3 \mu\text{M}$). The y-axis shows the difference between ΔAbs values at 443 nm and 416 nm from the difference spectra (Figure 4.18). The x-axis shows the concentration of NaCN. The data are fitted using the Hill equation (fit shown in blue) which produces a value of the apparent dissociation constant (K_H) of $540 \pm 20 \mu\text{M}$, and a Hill number of 3.00 ± 0.30 .

4.5. Electron Paramagnetic Resonance Spectroscopy

EPR was carried out on FLCYP116B1 and HDCYP116B1 to investigate the properties of the haem (as in Chapter 3 for HDCYP102A3), but also for the iron sulphur cluster [2Fe2S] in the reductase domain of FLCYP116B1. The low spin Fe^{3+} in the oxidised haem is paramagnetic and therefore EPR active. The [2Fe-2S] cluster is diamagnetic when oxidised due to spin pairing and therefore EPR silent. When reduced, the low-spin Fe^{2+} of the haem is diamagnetic and EPR silent. However, only one of the Fe atoms in the [2Fe-2S] cluster is reduced which breaks the spin coupling interaction making the cluster paramagnetic overall and therefore EPR active. For these reasons, spectra of the FLCYP116B1 and HDCYP116B1 were taken in the oxidised state and also under conditions of reduction by sodium dithionite.

Samples were prepared as previously described (see section 2.4.17.1) and the experiments were carried out by Dr. Stephen Rigby at the University of Manchester as described in section 2.4.17.3.

4.5.1. FLCYP116B1

The EPR spectra of oxidised and reduced FLCYP116B1 are presented in overview in Figure 4.20. As predicted, the oxidised spectrum shows signals corresponding to Fe^{3+} haem in a low-spin state while the [2Fe-2S] cluster is silent. In the reduced spectrum the [2Fe-2S] cluster is active and the haem is silent.

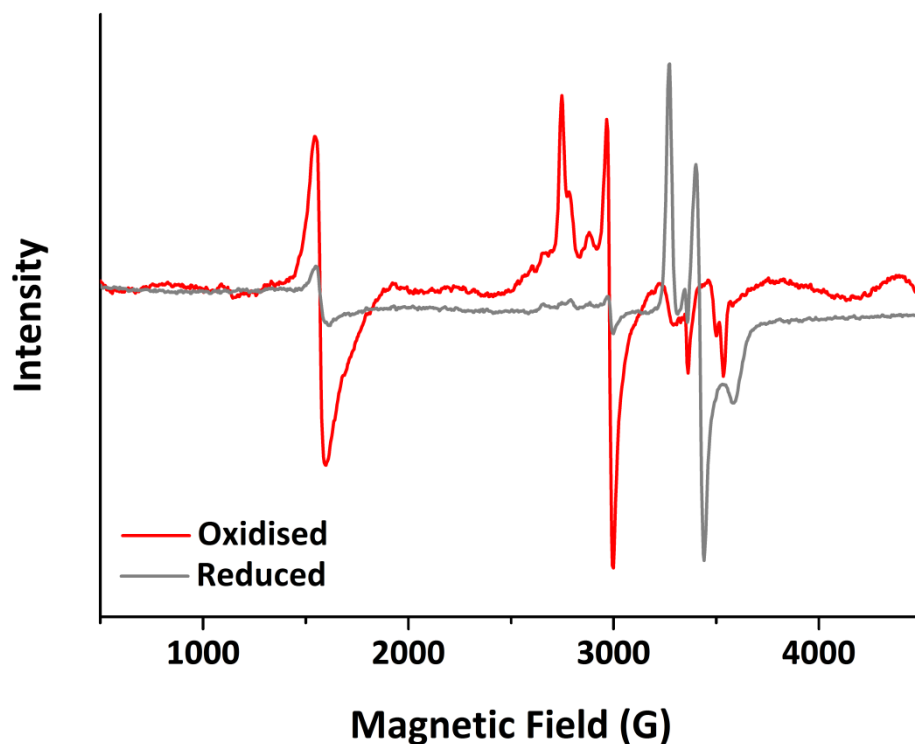


Figure 4.20. Electron paramagnetic resonance spectra of FLCYP116B1. The spectrum of the oxidised enzyme (200 μM) is shown in red and that of the dithionite reduced enzyme (200 μM) is shown in grey. In the oxidised spectrum the signals due to Fe^{3+} haem dominate the spectrum due to the [2Fe-2S] iron sulphur cluster in the reductase domain being EPR silent in its oxidised form. In the reduced spectrum, the haem signal disappears as Fe^{2+} is diamagnetic in its low-spin state and therefore EPR silent. Meanwhile, signals due to the [2Fe-2S] iron sulphur cluster (EPR active when reduced) become apparent.

An expansion of the EPR spectrum of oxidised FLCYP116B1 (Figure 4.21) shows the signals attributable to low-spin Fe^{3+} haem with g -values comparable to other P450 haems as shown for HDCYP102A3 (see section 3.2.3, Table 3.4). The g_x and g_z signals each appear split in two while the g_y signal is apparently unique; this is consistent with two populations of haem ligation. Ligation by imidazole during the purification protocol would seem to be a likely explanation, with one population of haem being imidazole ligated and the other being water (or HO^-) ligated. As demonstrated in section 5.2.2, imidazole binding to HDCYP116B1 haem seems to occur in two stages with the first showing tight binding ($K_d = 18 \pm 1 \mu\text{M}$) and the second showing weaker binding ($K_H = 580 \pm 5 \mu\text{M}$). It appears, from these spectra that the initial, more tightly bound, population of imidazole is retained in the purified FLCYP116B1, though this is not reflected in the UV/visible absorption spectrum. With the outer g_z component at 2.44, it is possible that bound imidazole is coordinated to the haem iron via its 6th water ligand to produce this set of g -values. It is interesting to note that the imidazole

remains bound despite the dilution involved in the dialysis and size exclusion chromatography steps in the purification (see section 2.4.8).

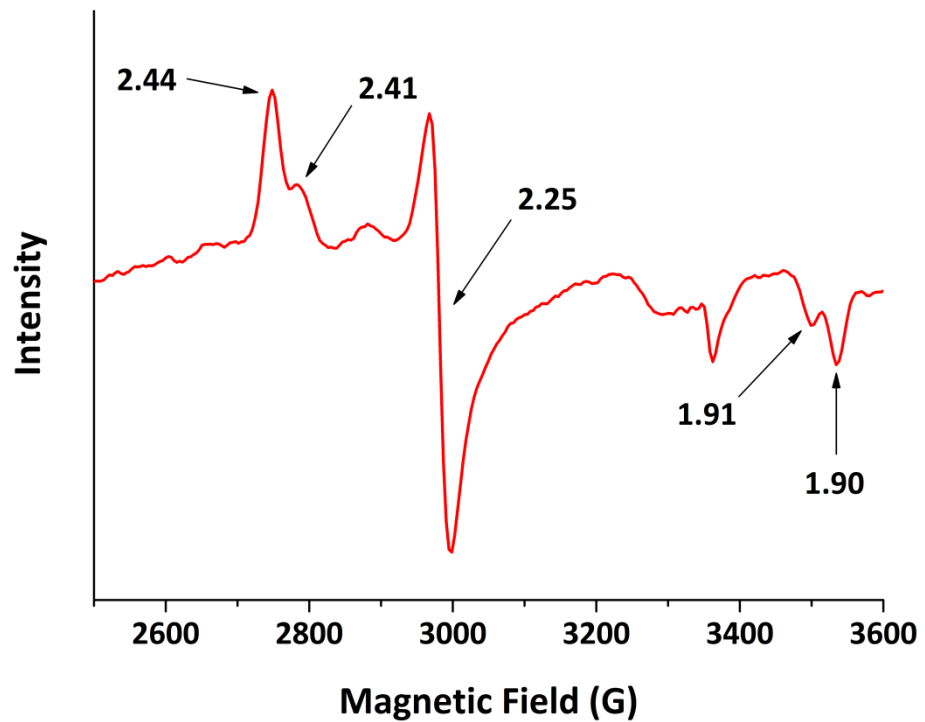


Figure 4.21. Expansion of the EPR spectrum (2500 – 3600 Gauss) of oxidised FLCYP116B1 (200 μ M) labelled with the corresponding g-values. $g_z = 2.44/2.41$, $g_y = 2.25$, $g_x = 1.91/1.90$. The splitting of the g_x and g_z signals could suggest that there are two populations of haem ligand states present in the FLCYP116B1 sample, although it may be likely that water (or hydroxide) is the 6th ligand in both species.

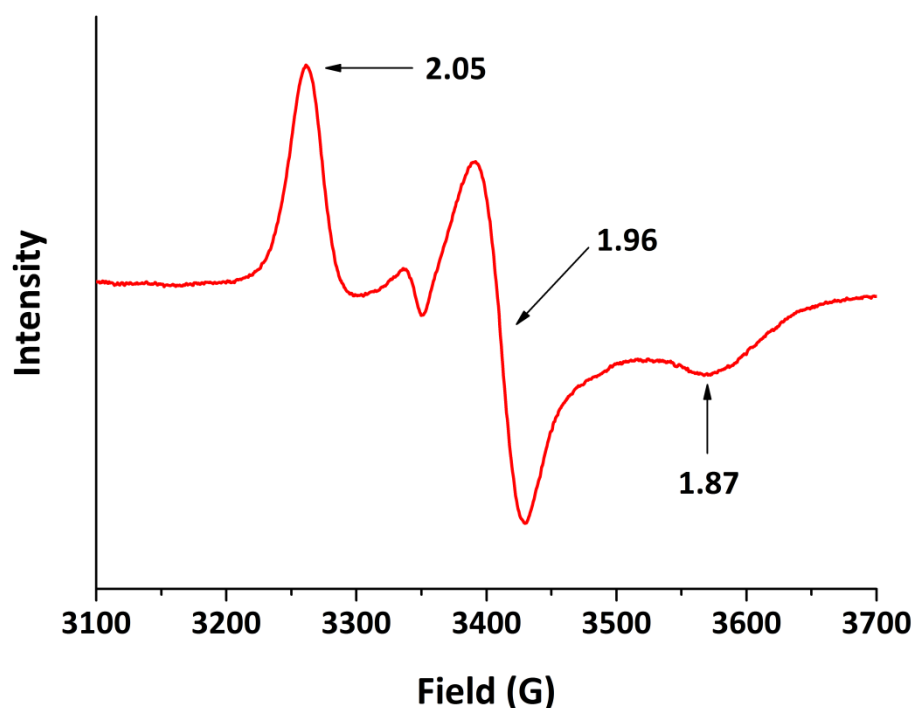


Figure 4.22. Expansion of the EPR spectrum of dithionite reduced FLCYP116B1 (200 μ M) labelled with the corresponding g -values ($g_z = 2.05$, $g_y = 1.96$ and $g_x = 1.87$). The signals in this region are due to the [2Fe-2S] cluster in the reductase domain.

An expansion of the dithionite reduced spectrum of FLCYP116B1 (Figure 4.22) shows signals that correspond fairly closely to previously measured signals for the reductase domain of CYP116B2[191] and for the two-electron reduced [2Fe-2S] cluster in PDOR from *Pseudomonas cepacia*[132] (see Table 4.4).

Enzyme	g_x	g_y	g_z
FLCYP116B1	1.87	1.96	2.05
RDCYP116B2	1.88	1.97	2.06
PDOR	1.90	1.95	2.04

Table 4.4. g -values from electron paramagnetic resonance spectroscopy studies of various enzymes related to CYP116B1. The characteristic signal, resulting from a reduced [2Fe-2S] cluster, consists of a rhombic trio of g -tensor elements designated x , y and z in accordance with axes of the molecular orbitals responsible. Data for other P450s are quoted from the literature. CYP116B2 (P450-Rhf) is the first P450-PDOR fusion discovered (discussed in section 4.1.2)[191]. PDOR is the homologous to the reductase domains of CYP116B2 and CYP116B1[132].

Integration of the EPR spectra showed that the ratio of [2Fe-2S] to haem present in the enzyme is 9.1:10 indicating that FLCYP116B1 is 91 % replete with [2Fe-2S] (making the reasonable assumption that the enzyme is fully replete with haem). This

quantitation was effected by double integration of the EPR signals using the software associated with the spectrometer operating system (Bruker, Coventry, UK).

Collectively, the EPR data for FLCYP116B1 are consistent with a “normal” thiolate-coordinated haem iron in the resting (oxidised) form of FLCYP116B1, although there is some heterogeneity in this species, likely resulting from some residual imidazole coordinating the haem iron (possibly indirectly, or else from aqua-ligated haem iron in different environments). The iron-sulfur and flavin centres are silent in the oxidised enzyme, but a strong signal consistent with a reduced [2Fe-2S] centre is seen in the dithionite reduced enzyme. There is no obvious radical signal in this state that could be associated with a flavin semiquinone, and thus it is likely that the flavin becomes 2-electron reduced and is converted to the EPR silent hydroquinone state. The haem is near fully reduced to an EPR silent ferrous state on dithionite reduction.

4.5.2. HDCYP116B1

As HDCYP116B1 could be purified without the use of imidazole (see section 4.3.1), EPR spectra were recorded to investigate the possible retained imidazole shown by the FLCYP116B1 EPR spectra.

The EPR spectrum of the imidazole-free HDCYP116B1 (Figure 4.23) did not show the splitting of the g_z and g_x signals shown in the spectrum of oxidised FLCYP116B1 (*c.f.* Figure 4.21).

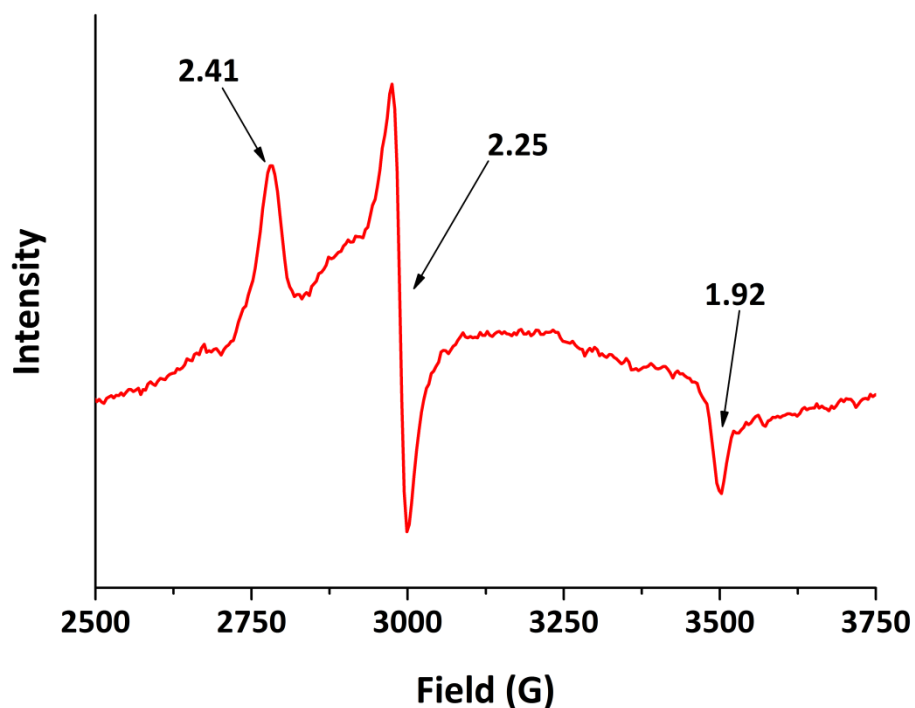


Figure 4.23. EPR spectrum of oxidised HDCYP116B1 (200 μ M) purified without imidazole. The spectrum shows the classic haem signal, a rhombic trio with g-values labelled on the spectrum ($g_z = 2.41$, $g_y = 2.25$ and $g_x = 1.92$). Splitting of the g_z and g_x signals, as shown for FLCYP116B1 purified in the presence of imidazole (see Figure 4.21), is not seen in this spectrum.

Next, imidazole was added to the HDCYP116B1 purified without imidazole and the EPR spectrum was recorded. This spectrum show splitting of g_z and g_x into three signals each. Two each of these signals correspond closely to those found in the FLCYP116B1 oxidised EPR spectrum. This strongly supports the suggestion that the splitting of signals is due to the retention of imidazole from the purification process of the enzyme. Based on the EPR spectrum for FLCYP116B1, the sets of g-values may be assigned to imidazole bound directly to the haem iron as a 6th ligand (2.53, 2.25, 1.86), imidazole bound indirectly via water, retained as the 6th ligand (2.44, 2.25, 1.89) and a residual amount of P450 with no imidazole bound and the 6th water ligand in place (2.41, 2.25, 1.92).

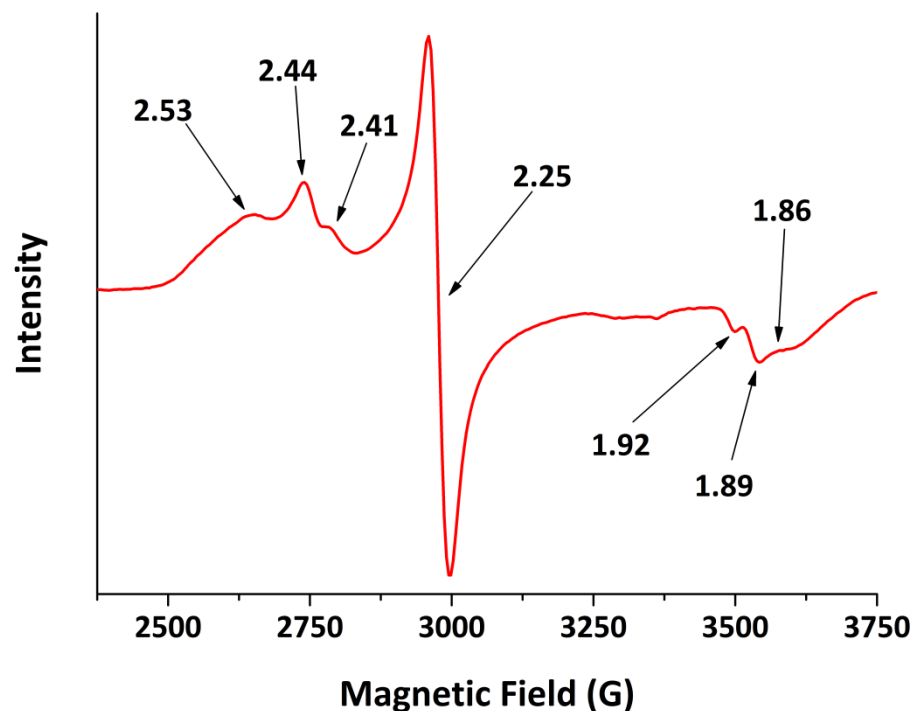


Figure 4.24. EPR spectrum of oxidised HDCYP116B1 (200 μ M) purified without imidazole, and with 1 mM imidazole added. The spectrum shows the classic haem signal, a rhombic trio with g-values labelled on the spectrum. The g_z and g_x signals are each split into three, with two of each having g-values corresponding to those found in the FLCYP116B1 oxidised spectrum (see Figure 4.21).

4.6. Differential Scanning Calorimetry

DSC was carried out on the HDCYP116B1 and FLCYP116B1 proteins to assess the relative stability of the intact fusion enzyme compared to the isolated haem domain and, indeed, the intrinsic stability of both to thermal denaturation.

4.6.1. HDCYP116B1

DSC experiments were carried out on HDCYP116B1 protein in the presence and absence of the ligand 2-phenylimidazole. In both cases the data recorded were of a reasonable quality, and certainly better than that acquired for HDCYP102A3 (see section 3.2.6). The heat capacity profiles showed what appeared to be two unfolding events in the region 35 – 50 $^{\circ}$ C. Curve-fitting allowed individual unfolding events and associated T_m values to be separated from the original data.

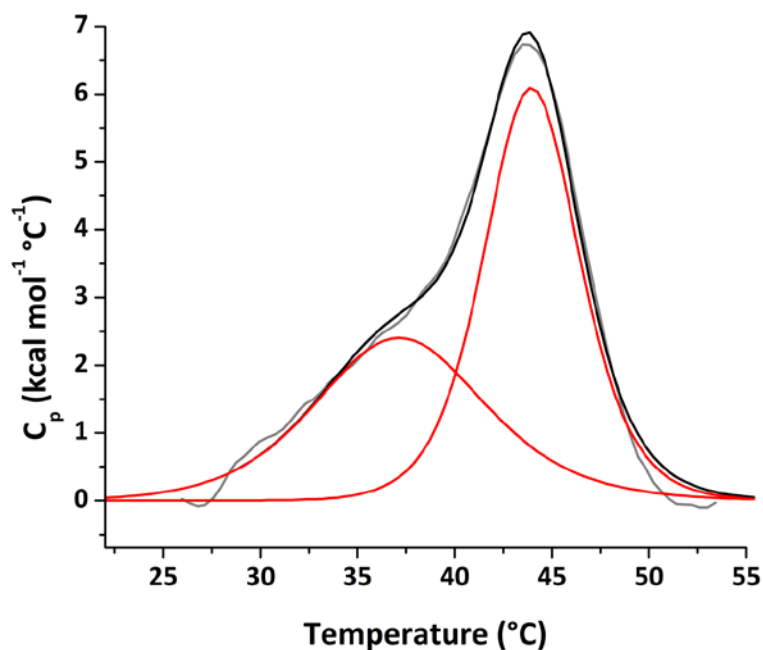


Figure 4.25. Plot of results of a DSC experiment on HDCYP116B1 ($\sim 20 \mu\text{M}$). The heat capacity profile (grey line) was curve-fitted to derive the individual unfolding events making a contribution. Two curves were fitted and are shown (red lines) along with the sum of the fitted curves (black line) which allows comparison of the fit to the original data. T_m values associated with the two unfolding events represented by the two curves are $37.1 \text{ }^\circ\text{C}$ (T_{m1}) and $43.9 \text{ }^\circ\text{C}$ (T_{m2}).

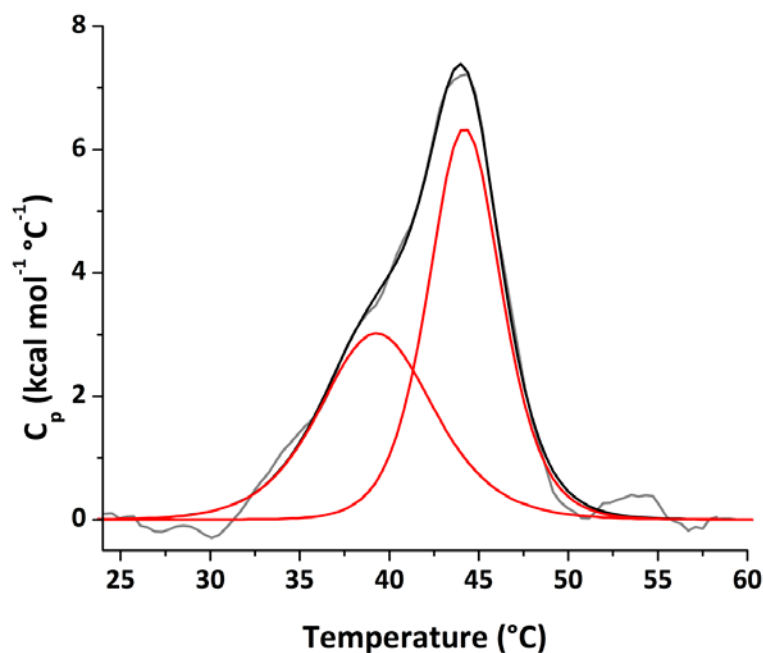


Figure 4.26. Plot of the results of a DSC experiment on HDCYP116B1 ($\sim 20 \mu\text{M}$) in the presence of 2-phenylimidazole (1 mM). The heat capacity profile (grey line) was curve-fitted to derive the individual unfolding events making a contribution. Two curves were fitted and are shown (red lines) along with the sum of the fitted curves (black line) which allows comparison of the fit to the original data. T_m values associated with the two unfolding events represented by the two curves are $39.2 \text{ }^\circ\text{C}$ (T_{m1}) and $44.4 \text{ }^\circ\text{C}$ (T_{m2}).

For HDCYP116B1 in the absence of ligand the values were $T_{m1} = 37.1$ °C and $T_{m2} = 43.9$ °C while in the presence of 2-phenylimidazole there was an increase in both values: $T_{m1} = 39.2$ °C and $T_{m2} = 44.4$ °C. This indicates, therefore, that the binding of 2-phenylimidazole has a stabilising effect upon the HDCYP116B1 protein, shifting T_{m1} by 2.1 °C and T_{m2} by 0.5 °C. The data here are also consistent with the previous data set for the HDCYP102A3, in terms of the likelihood here (despite the eventual aggregation and poorer data set for this protein) that this protein also has a 2-stage thermal unfolding profile. Potentially, for both HDCYP102A3 and HDCYP116B1, this reflects the consecutive unfolding of the smaller beta domains and the larger alpha domains of the proteins.

4.6.2. FLCYP116B1

The experiments performed above for HDCYP116B1 were repeated for FLCYP116B1. In both the 2-phenylimidazole-bound and ligand-free runs, two unfolding events were again observed. This may be because any changes in C_p associated with the reductase domain overlap with those for the haem domain. Certainly the peak in C_p is, in both cases, broader than that observed for HDCYP116B1. For ligand-free FLCYP116B1, T_m values were 37.9 °C and 41.9 °C, while for the 2-phenylimidazole-bound FLCYP116B1 T_m values were 40.9 °C and 49.8 °C.

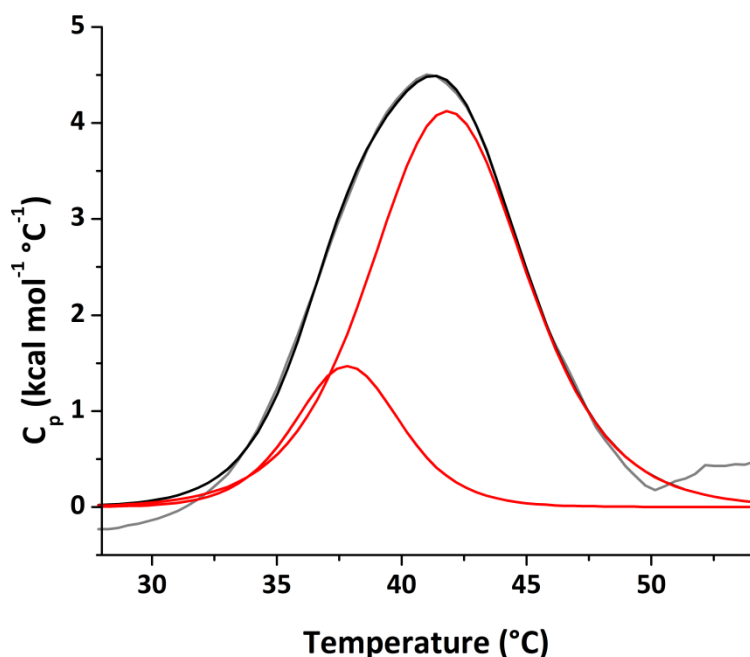


Figure 4.27. Plot of the results of a DSC experiment on FLCYP116B1 ($\sim 20 \mu\text{M}$). The heat capacity profile (grey line) was curve-fitted to derive the individual unfolding events making a contribution. Two curves were fitted and are shown (red lines) along with the sum of the fitted curves (black line) which allows comparison of the fit to the original data. T_m values associated with the two unfolding events represented by the two curves are $37.9 \text{ }^\circ\text{C}$ (T_{m1}) and $41.9 \text{ }^\circ\text{C}$ (T_{m2}).

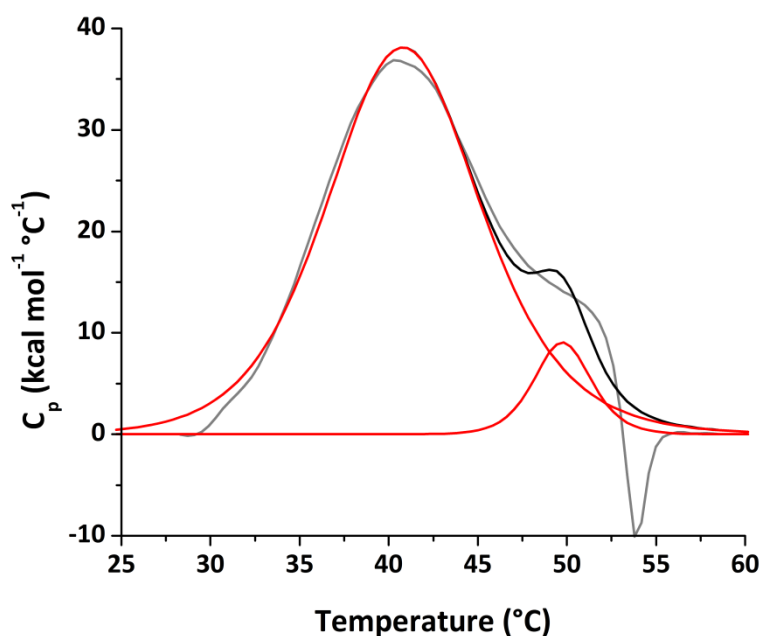


Figure 4.28. Plot of the results of a DSC experiment on FLCYP116B1 ($\sim 20 \mu\text{M}$) in the presence of 2-phenylimidazole (1 mM). The heat capacity profile (grey line) was curve-fitted to derive the individual unfolding events making a contribution. Two curves were fitted and are shown (red lines) along with the sum of the fitted curves (black line) which allows comparison of the fit to the original data. T_m values associated with the two unfolding events represented by the two curves are $40.9 \text{ }^\circ\text{C}$ (T_{m1}) and $49.8 \text{ }^\circ\text{C}$ (T_{m2}).

For both FLCYP116B1 and HDCYP116B1, the T_m values, determined by DSC point to a stabilisation in the inhibitor-bound forms. This may be attributable to a conformational change, effected by ligand binding, leading to enhanced stability of the proteins. It should also be noted, as for HDCYP102A3, that all T_m values are above the normal temperature range experienced by the CYP116B1 host organism *C. metallidurans*, a soil bacterium of the order *Burkholderiales*. Thus, CYP116B1 proteins are likely quite stable to temperature conditions normally encountered by the bacterium.

4.7. Multi-Angle Light Scattering

The HDCYP116B1 and FLCYP116B1 proteins purified were analysed by MALS in the same way as for HDCYP102A3 (described in 3.2.8). The data collected were fitted by the Zimm fitting method[136, 173] to derive values for the molecular mass and hydrodynamic radius of species present in each sample.

4.7.1. FLCYP116B1

MALS was carried out on the FLCYP116B1 protein to assess the oligomeric state and homogeneity. As stated in chapter 3, the prototype P450-reductase fusion, BM3 has been shown to exist in solution as a dimer and the dimeric state appears critical to the catalytic function of the enzyme as a fatty acid hydroxylase[73]. However, it has also been shown that CYP116B2 (P450-RhF) most likely exists in solution and is catalytically active as a monomer[191]. It might therefore be expected that FLCYP116B1 should also exist as a monomer given its relation with CYP116B2.

The MALS experiment showed three distinct species present in the purified FLCYP116B1 sample: one major species and two minor species eluting from the size exclusion column either side of the major peak (Figure 4.29).

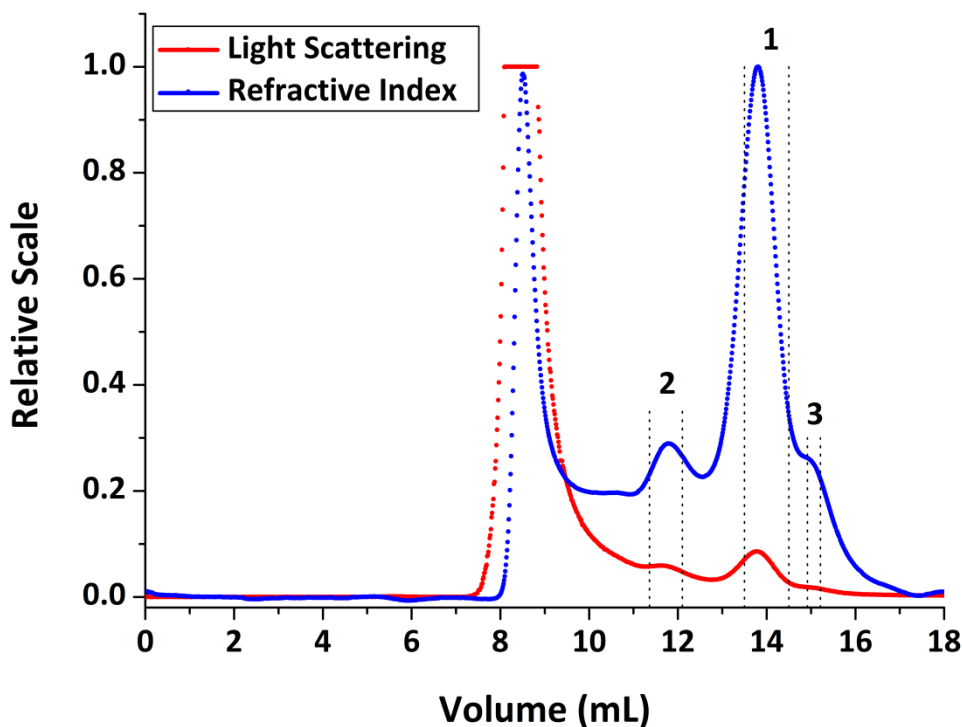


Figure 4.29. Trace from a MALS run of FLCYP116B1, showing light scattering and refractive index against elution volume from a size exclusion column. Three peaks are labelled, indicating three distinct species present in the sample.

The major species (denoted peak 1 in all relevant figures) showed an average molecular mass of 93 ± 0.1 kDa and hydrodynamic radius of 3.8 ± 0.03 nm. The value of the molecular mass is close to that estimated for the FLCYP116B1 protein (86 kDa) and the difference can be explained as the value calculated from MALS is based on the hydrodynamic radius. This peak would therefore seem to represent monomeric FLCYP116B1.

The first of the minor species (denoted peak 2 in relevant figures) showed an average molecular mass of 226 ± 0.2 kDa and hydrodynamic radius of 5.7 ± 0.05 nm. This species is likely to represent dimeric FLCYP116B1 as a minor component of the sample.

The second minor peak (denoted peak 3 in relevant figures) showed an average molecular mass of 78 ± 0.5 kDa and a hydrodynamic radius of 2.9 ± 0.1 nm. This peak could be due to a minor contaminant in the sample, although FLCYP116B1 purity was confirmed by SDS PAGE as shown above. It could also potentially be an alternative conformation of the FLCYP116B1 protein or a partially proteolysed or degraded species.

For each peak the values of molecular mass and hydrodynamic radius recorded during the experiment are plotted in Figure 4.30. For the major species (peak 1) the values are the most consistent. The first minor species (peak 2) shows more variation in these values across the peak. The second minor species (peak 3) show a fairly consistent molecular mass, but a large variation in hydrodynamic radius.

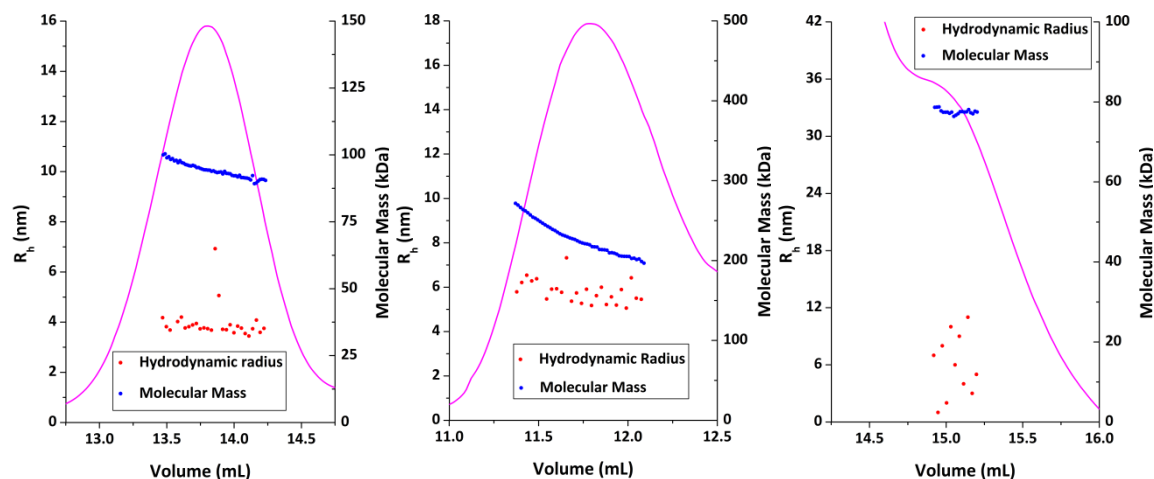


Figure 4.30. Plots of values of hydrodynamic radius (R_h) and molecular mass for each of the three peaks in the MALS trace for FLCYP116B1 (Figure 4.29) (shown in magenta): Peak 1 (left), peak 2 (centre) and peak 3 (right).

4.7.2. HDCYP116B1

For HDCYP116B1 the more immediate goal of structural characterisation by X-ray crystallography demanded that the monodispersity of the sample be examined by MALS to insure that it would be suitable for crystallisation.

The MALS experiment showed a single species present in the HDCYP116B1 sample (Figure 4.31). This species showed an average molecular mass of 59.9 ± 0.2 kDa and a hydrodynamic radius of 3.1 ± 0.03 nm. The value of molecular mass is close to that estimated for the HDCYP116B1 protein (50.7 kDa) and the hydrodynamic radius indicates a compact protein.

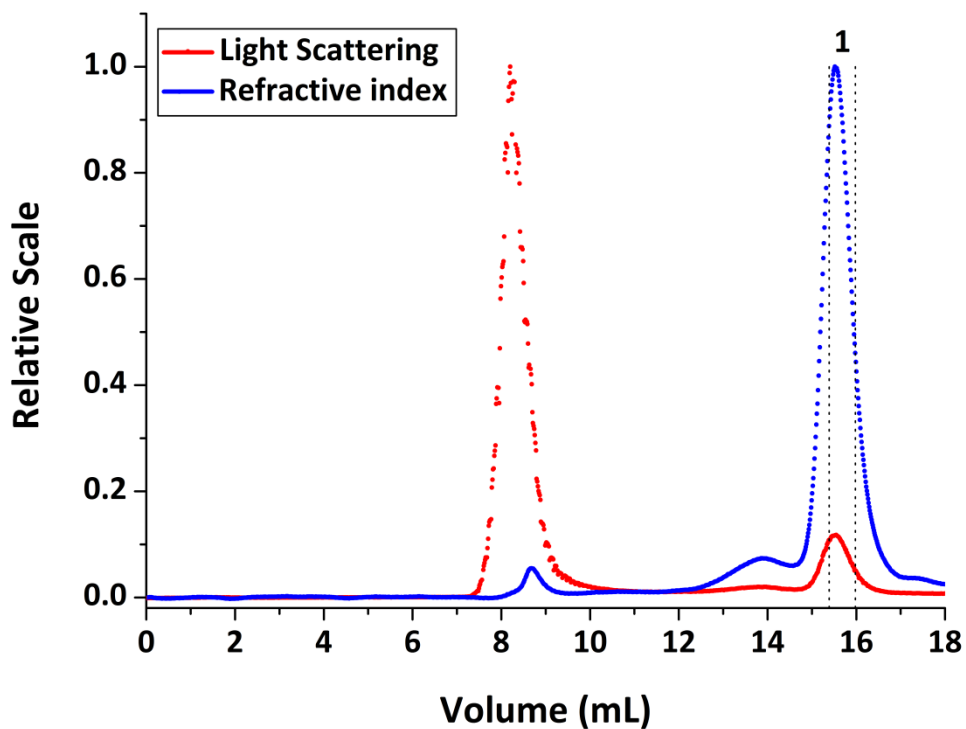


Figure 4.31. Trace from a MALS run of HDCYP116B1, showing light scattering and refractive index against elution volume from a size exclusion column. A single peak is labelled, indicating a homogenous sample of one protein species.

The values of molecular mass and hydrodynamic radius recorded during the experiment are shown in Figure 4.32. This plot shows a low level of variance of both values across the peak.

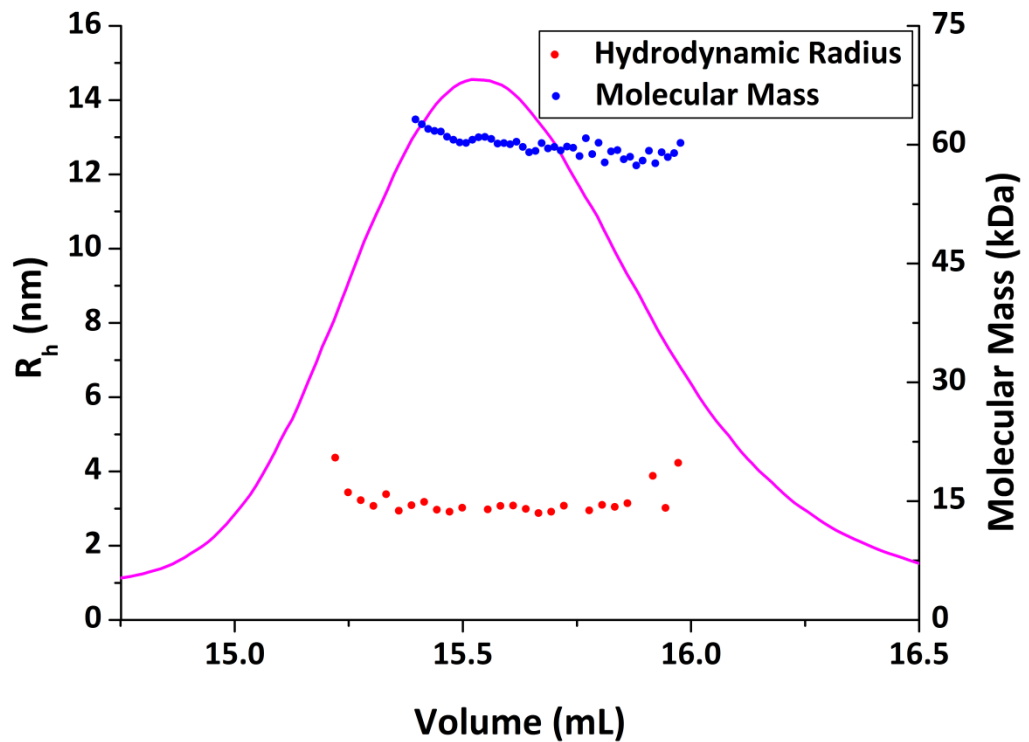


Figure 4.32. Plots of values of hydrodynamic radius (R_h) and molecular mass for the peak in the MALS trace of HDCYP116B1 (Figure 4.31) (shown in magenta).

Conclusions from the MALS data are that there is potentially a small amount of the dimeric form of FLCYP116B1 present in solution, but that the protein is mainly monomeric. This situation is different from that of the catalytically self-sufficient BM3 (CYP102A1) enzyme, which is functional in substrate oxidation in the dimeric form[73]. The corollary of this finding is likely that electron transfer in the CYP116B1 (and other family members) occurs within a monomer, as opposed to between monomers in a dimer, as is seen in BM3 and the related nitric oxide synthase flavocytochromes[73, 195]

4.8. Crystallographic Trials

Crystallographic trials were carried out upon the HDCYP116B1 protein. The pure protein sample, with its homogeneity assured by light scattering experiments was diluted to concentrations of 10 and 12 mg mL⁻¹. For some P450 enzymes crystallisation has only been possible with a ligand bound form of the enzyme. To this end 10 and 12 mg mL⁻¹ stocks were also produced with the addition of 720 µM 2-phenylimidazole. These protein stocks were used with crystallographic screens Classics, Classics II, PEGs, Anions, JCSG + and JCSG II suites. Screening trays were set up using the Mosquito robot as described in 2.4.10 with 200 nL drops (100 + 100).

No crystals were produced in any of the screening conditions trialled. As the homogeneity of the protein stock solution had been established by MALS experiments there is no obvious explanation for the non-appearance of crystals, aside from no appropriate conditions for crystallogenesis having been found.

4.9. Homology Modelling

In the absence of a crystal structure, a structural model of the HDCYP116B1 protein was produced using Swiss-Model, an online, fully automated homology modelling tool. As described in 3.2.10 this tool chooses a template structure from the PDB using sequence alignment and implied structural similarity. The limitation of this technique is that the resultant structure may not represent the actual conformation of the protein *in vivo* as this is implied from the template. The P450 haem and reductase domains of CYP116B1 were modelled separately against two different template structures.

4.9.1. HDCYP116B1

The haem domain model is based on the structure of CYP105P1, from *Streptomyces avermitilis*, in complex with filipin I (an antibiotic) (RCSB pdb accession code: 3ABA). This enzyme shows 27 % sequence identity with HDCYP116B1.

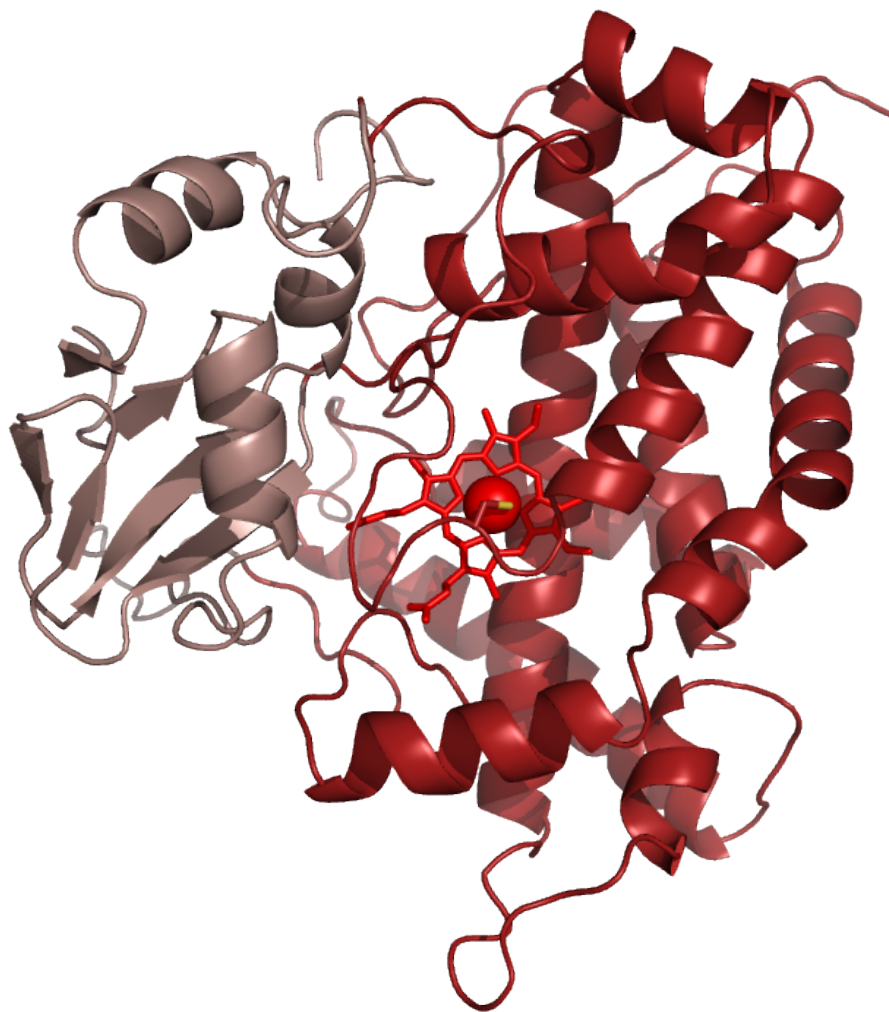


Figure 4.33 Image of a homology model of the haem domain of CYP116B1, shown as a cartoon. The α and β subdomains are shown in dark red and salmon colours respectively. The structure and position of the haem cofactor is shown in red and the side-chain of the ligating cysteine residue (Cys-385) is also shown. Behind the haem and spanning the length of the α subdomain is the I-helix.

The model shows a globular looking structure and displays recognisable elements of the P450 fold. The α and β subdomains, the long I-helix and the haem-ligating cysteine residue are all visible and sensibly aligned (see Figure 4.33). Also present is Thr-278, in a position on the I-helix analogous to Thr-268 in BM3 (see Figure 4.34), which is likely involved in interactions with, and/or proton transfer to, dioxygen-bound form(s) of the haem iron during the catalytic of fatty acid hydroxylation[196].

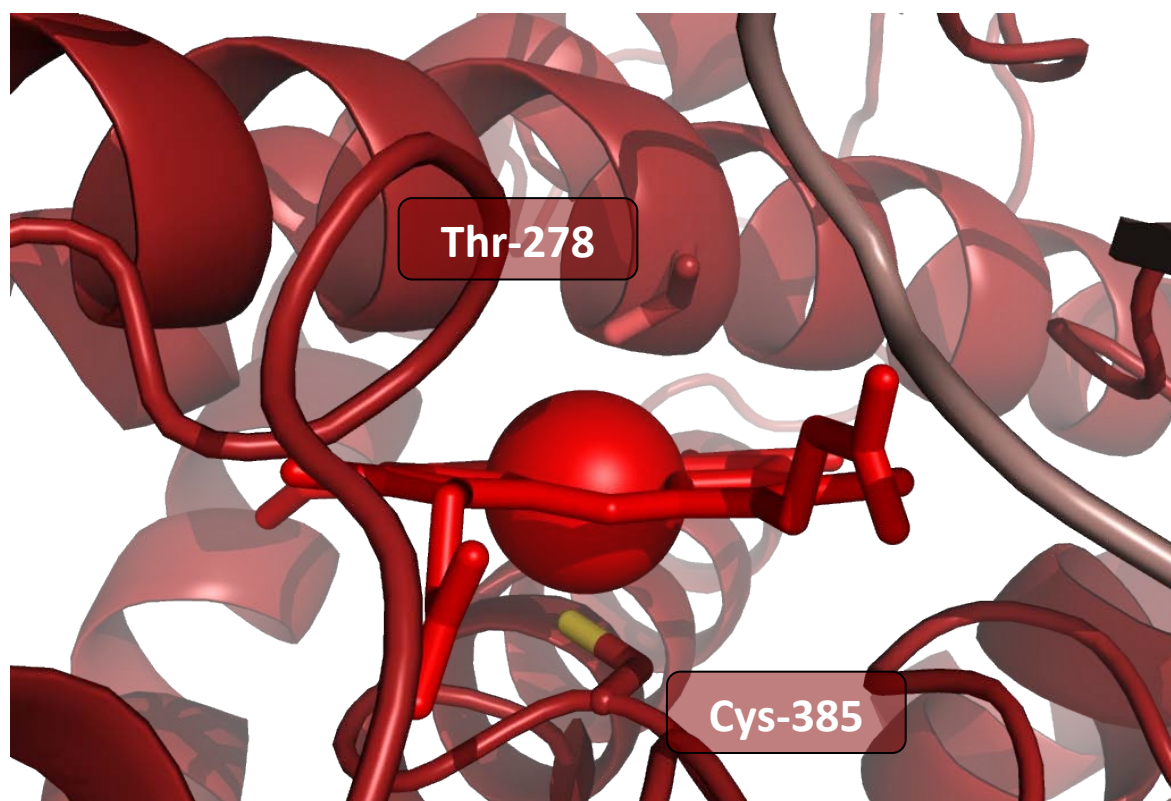


Figure 4.34 Image of a homology model of the haem domain of CYP116B1, shown as a cartoon. The view between the haem cofactor and the I-helix is shown in this image. The ligating cysteine (Cys-385) can be seen below the haem, and on the distal side is seen Thr-278 which is predicted to exert control over the oxygen activation and substrate hydroxylation steps by analogy to Thr-268 in BM3[196]

Despite these familiar elements, the model does display a large number of loops compared to other P450 haem domain structures. This may be a reflection on the secondary structural content or assignment in the template structure, or else on the difficulty in producing a convincing model with low sequence identity.

The cavity around the haem in this model contains many hydrophobic alkyl side-chains (*i.e.* valine, alanine, leucine *etc.*) and there are several aromatic and polar residues, further from the haem, which might mediate binding interactions with different moieties of substrate molecules. These speculations must, again, be subject to the caveat that the model may not represent a highly accurate representation of the structure of HDCYP116B1.

4.9.2. RDCYP116B1

The model of the reductase domain also shares a low sequence identity (32 %) with its template structure, the structure of phthalate dioxygenase reductase from

Pseudomonas cepacia (RCSB PDB accession code: 2PIA). In this model, again, the two domains (Fd-like and FdR-like) are clearly visible, with the FMN- and NAD(P)H-binding subdomains of the FdR-like domain also distinguishable (see Figure 4.35).

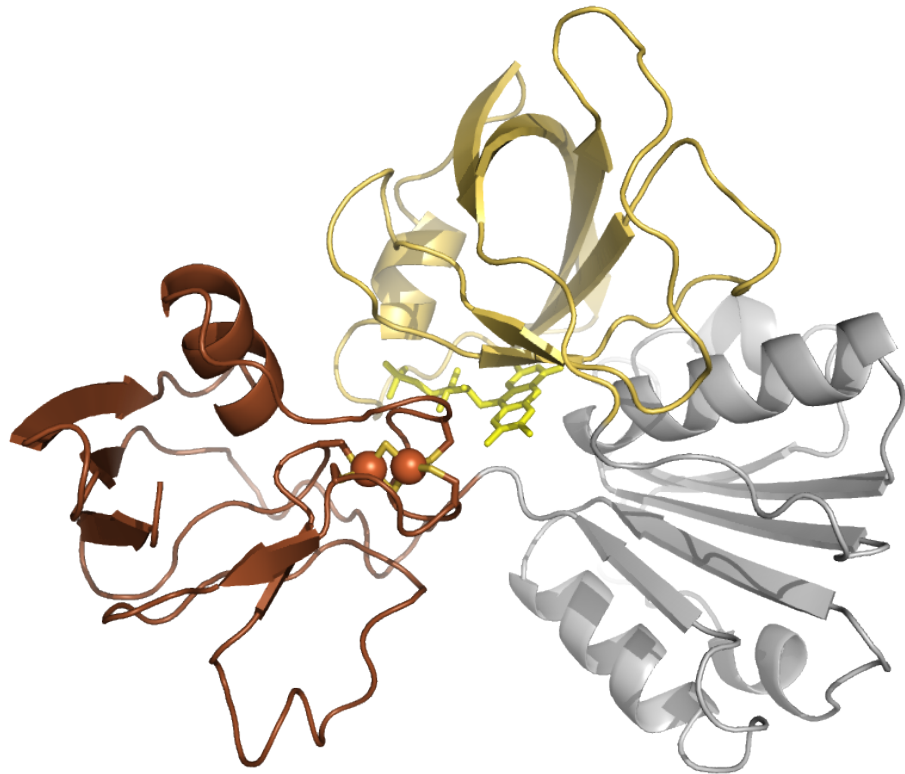


Figure 4.35 Image of a homology model of RDCYP116B1 shown as a cartoon. The template structure is that of phthalate dioxygenase reductase from *Pseudomonas cepacia* (RCSB PDB accession code: 2PIA). The Fd-like domain is coloured brown, while the FMN- and NAD(P)H-binding subdomains of the FdR-like domain are coloured gold and silver respectively. The cofactors within these subdomains, [2Fe-2S] cluster (orange) and FMN (yellow) are also shown.

Binding sequences for the two redox cofactors present in RDCYP116B1 are present in the model (see Figure 4.36). The ribityl chain of FMN in PDOR is shown to interact with Arg-80 in the sequence GRGGS[132], and the equivalent sequence in RDCYP116B1 is SRGGS with Arg-83 being the interacting residue. The presence of the additional serine residue (Ser-82) may provide additional binding interactions with this part of the cofactor. Phe-224 (Phe-225 in PDOR) is thought to interact with both the flavin isoalloxazine ring and the nicotinamide ring of NAD(P)H and to mediate electron transfer between the two[132]. The [2Fe-2S] is a plant-type iron sulfur cluster and, as such, is ligated by four cysteine residues (Cys-273, -276, -281 and -311).

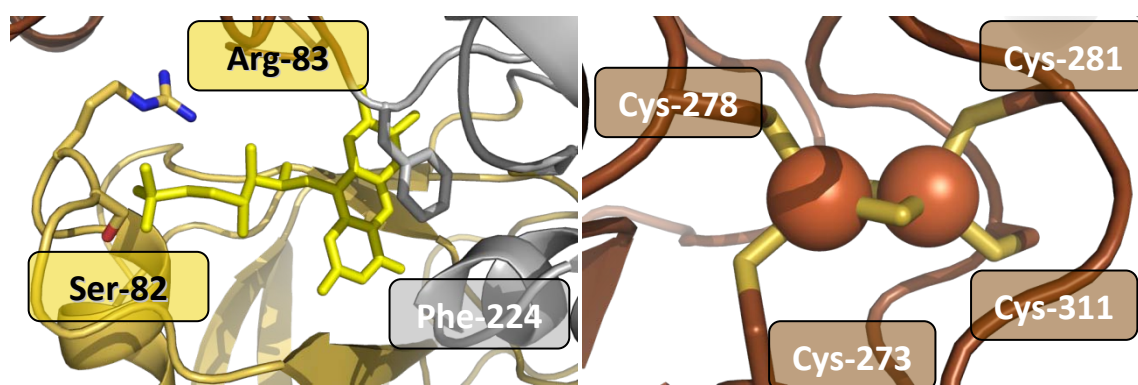


Figure 4.36. Images of the homology model of the reductase domain of CYP116B1, focussing on the redox cofactors. The left image shows the FMN and Arg-83, analogous to Arg-80 in PDOR, shown to interact with the ribityl chain of the FMN[132]. This image also shows Ser-82, (the equivalent position in PDOR is occupied by glycine) and Phe-224. The right image shows the [2Fe-2S] cluster and the ligating cysteine residues (273, 276, 281 and 311).

4.10. Summary & Discussion

In this chapter data are presented that cover the study of CYP116B1 and its isolated domains in the context of their spectroscopic, biophysical and structural properties. Initially, cloning work was carried out to optimise the expression and purification of FLCYP116B1 and HDCYP116B1. This involved employing a 6xHis tag vector and correcting an error in the gene sequence which would have adversely affected expression of CYP116B1 proteins. This work successfully improved the yield of these proteins over previous schemes, and greatly simplified the purification regime. The disadvantage to these new purification protocols was the generation of a mixed ligation state of the haem cofactor of the proteins. This issue was discovered to be connected to an unusual binding profile for the haem ligand imidazole with CYP116B1 proteins. This issue was, at least partially, overcome with an imidazole-free purification protocol for HDCYP116B1.

The CYP116B1 reductase domain also required work to correct a random nucleotide deletion which would have adversely affected its expression. A purification scheme was developed for this domain which was successful, if not entirely reproducible.

The P450 haem cofactors in the FLCYP116B1 and HDCYP116B1 proteins were characterised by the formation of the CN-bound and reduced CO-bound complexes. The dithionite-reduced CO-bound spectra of FLCYP116B1 and HDCYP116B1 each showed two populations of haem Soret bands corresponding to P450 and P420 complexes. This reveals that in a population of CYP116B1 the thiolate group of the haem-ligating cysteine residue is protonated to a thiol upon reduction of the haem by dithionite. This distinguishes CYP116B1 from CYP102A1 and related enzymes, which mainly retain thiolate ligation upon reduction in this manner. Studies on other P450s show that thiolate protonation is common in reduced substrate-free enzymes, but that substrate binding can stabilise the thiolate ligand in the reduced form[197, 198].

EPR spectroscopy was used to further characterise the redox cofactors in the HDCYP116B1 and FLCYP116B1 proteins. The spectra of the full-length protein and isolated P450 domain, in their oxidised state, showed signals similar to those recorded for other P450 haems. Additionally, these spectra confirmed that the purification of the FLCYP116B1 protein had left residual imidazole bound to its haem cofactor. The EPR spectrum of FLCYP116B1 in a dithionite-reduced state revealed a signal due to the reduced [2Fe-2S] cluster. This signal was comparable to that

recorded for the related protein phthalate dioxygenase reductase[132]. The lack of any signal for a flavin radical indicates that the FMN cofactor is reduced directly to the hydroquinone form.

The thermal stability of the FLCYP116B1 and HDCYP116B1 proteins was investigated by DSC and the stabilising effect of a haem ligand, 2-phenylimidazole, was examined. These experiments revealed a profile of two unfolding events for both FLCYP116B1 and HDCYP116B1, although, in the former, unfolding events arising from the reductase domain may overlay with those of the haem domain and appear as one broader event. Shifts in T_m values, for the unfolding events, show that the binding of the ligand 2-phenylimidazole has a stabilising effect (with regard to temperature) over both the haem domain and full-length protein. Additionally, these values suggest that the HDCYP116B1 and FLCYP116B1 proteins are stable at temperatures up to 37 °C, well above what the host organism *C. metallidurans* (a soil bacterium) would likely encounter.

A MALS experiment was carried out on HDCYP116B1 prior to an attempt to crystallise the protein to attempt structural characterisation. This experiment showed that the HDCYP116B1 protein sample purified was entirely monodisperse, with one species, a monomer, present in solution.

The situation for the FLCYP116B1 protein was shown not to be so clear-cut when MALS was carried out. Three distinct species were observed, the predominant of which was a monomer, with two minor species also present. The first had an implied molecular weight corresponding closely to an FLCYP116B1 dimer. Work on CYP116B2 has shown that it is catalytically active as a monomer and it would be expected that the same would be true of CYP116B1. It is possible that the dimer only occurs at the relatively high concentrations (compared to *in vivo*) in which it was purified. The third species appeared to be a partly proteolysed or an alternate conformation of FLCYP116B1, with a lower implied molecular weight than the monomer.

As with the work on HDCYP102A3 presented in chapter 3, structural characterisation of HDCYP116B1 by x-ray crystallography was an aim in this study. Crystal trials were conducted in the same manner, but produced no crystals. Once again, to partly compensate for the lack of structural characterisation, homology modelling was

carried out. Due to the lack of a close homologue with structural characterisation, the results for HDCYP116B1 are less reliable than those for HDCYP102A3. Nevertheless, key features of the P450 haem domain can be seen in the model and even some of the key residues in the haem binding cavity.

The RDCYP116B1 protein was also modelled, based upon the structure of PDOR, an enzyme thought to evolutionarily related. The model showed that the RDCYP116B1 protein shares the same domain structure as PDOR with [2Fe-2S], FMN and NAD(P)H-binding subdomains, and key residues involved in FMN interactions and iron atom ligation identified.

Following the work presented in this chapter, further attempts at structural characterisation of the isolated domains of CYP116B1 by x-ray crystallography could be undertaken. Aside from the failure to produce HDCYP116B1 crystals under the screening conditions applied, there is nothing to discourage this approach. Expanded trials including using bound ligands (*e.g.* 2-phenylimidazole) might yet yield viable crystals. In addition, if a more reliable purification scheme could be developed for the RDCYP116B1 protein, this might also present a target for structural characterisation to further probe any relevant variations from the structure of PDOR. One possible means of improving the purification might be a similar approach as was taken with the other CYP116B1 proteins: the use of a 6xHis tag to enable purification using affinity chromatography. Further work on CYP116B1 could provide valuable insights into the structure/mechanism relationship in this P450-reductase fusion.

Additionally, structural aspects of the CYP116B1 proteins could be probed using other techniques such as CD or, as demonstrated in the work in chapter 3, resonance Raman spectroscopy to investigate the haem structure in its resting state and when bound to inhibitory ligands or substrates such as thiocarbamate herbicides (see Chapter 5)

As shown in the first section of this chapter there are many other potential P450 fusions apparently fulfilling diverse roles in a variety of organisms. These include possible novel redox partner systems, enzymes with related functions to P450s and a range of other functional domains fused with P450 enzymes in a single polypeptide chain.

As whole-genome sequencing projects continue to amass data, the number and variety of fusions discovered will inevitably increase. As it stands only three of the P450-fusions presented have been confirmed and characterised. Thus, there remains immense scope for further research in the area of P450 fusion enzymes.

Indeed, a recently discovered example of a P450 fusion with an enzyme of related function is the P450-peroxidase (haem dioxygenase) fusion found in fungal species of the genus *Aspergillus*[199]. These enzymes, known by the abbreviation Ppo (psi-factor producing oxygenases), perform consecutive reactions upon fatty acid substrates to produce psi (precocious sexual indicator) factors[200]. This demonstrates the continuing research, and growing interest in the area of P450 fusions.

Characterisation of Novel Cytochrome P450 Fusion Systems

Chapter 5

CYP116B1 Enzyme Studies

5. CYP116B1 Enzyme Studies

5.1. General Information

Though the physiological role of CYP116B1 is not clear, two compounds have been identified that are bound in the active site and are hydroxylated under conditions of turnover (work by Warman *et al.* [135]). These substrates, EPTC and vernolate, are members of the thiocarbamate class of organic compounds and are both agricultural herbicides[201]. As discussed in the previous chapter (section 4.1.2), related enzymes have been found to have various roles in the degradation of thiocarbamates and other pesticides in species of soil bacteria able to utilise them as a carbon source. For example, CYP116A1 from *Rhodococcus sp.* strain NI86/21 which confers upon this organism the ability to degrade EPTC and atrazine[118]. In this case the proposed role of the enzyme is hydroxylation of EPTC as an activation step leading to dealkylation of the molecule[119]. In the proposed mechanism, CYP116A1 hydroxylates EPTC at the α carbon of one of the *N*-propyl groups. The hydroxy EPTC product decays leading to dealkylation and the release of propanal. An aldehyde dehydrogenase, shown to be induced along with CYP116A1 by EPTC, then converts the propanal into propanoic acid[202] (see Figure 5.1).

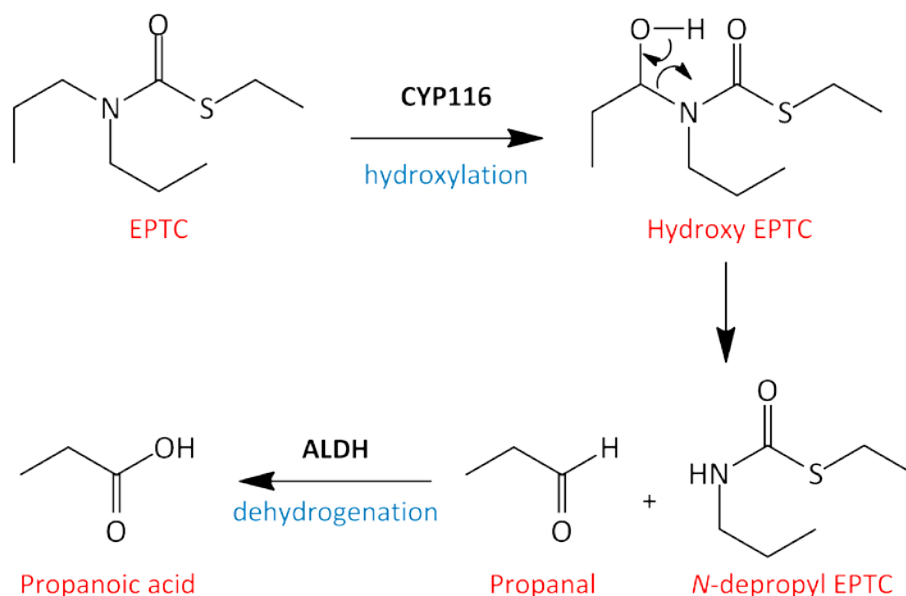


Figure 5.1. Proposed mechanism of the degradation of EPTC involving CYP116 enzymes. The first step is hydroxylation at the α carbon of one of the *N*-propyl groups to produce hydroxy EPTC. This species then decays to leave *N*-depropyl EPTC and propanal. The second enzymatic step involves the conversion of the propanal to a carboxylic acid by an aldehyde dehydrogenase (ALDH).

This mechanism of activity is supported by work showing *N*-depropyl EPTC as a metabolite of the reaction with CYP116A1[119]. If similar reactivity could be shown and/or developed in CYP116B1, its fused nature would make it a convenient single unit that could be engineered into other organisms for the purposes of bioremediation. This is the most immediate potential biotechnological application and if successful others might follow in fields such as synthetic chemistry.

5.1.1. Bioremediation with P450s

There is existing work towards bioremediation applications using P450 enzymes including CYP177A1 (XplA) from *Rhodococcus rhodochrous*, mentioned previously in section 4.1.1.1. This enzyme is a fusion of P450 and flavodoxin reductase domains which employs a separate redox partner (XplB) resembling flavodoxin reductase[180]. The P450 system is capable of degrading the explosive compound RDX with a proposed mechanism of two denitration steps followed by ring cleavage[203]. Investigations have been made of a possible application for treating soil contaminated with RDX using the XplA-B system[204].

Approaches using BM3 are mainly focussed at taking the robust electron transfer system of the fusion enzyme system and developing, using mutagenic techniques, new activities which could be useful for bioremediation or for generation of high value molecules such as drug metabolites. A general discussion of work aimed at altering the substrate specificity of BM3 is presented in section 1.1.10.

The work described in this chapter provides initial steps towards a greater understanding the reactivity and substrate specificity of CYP116B1. This may inform future work on this enzyme aimed at bioremediation applications.

5.1.2. Experimental Outline

The experimental aim of the work in this chapter was to identify some characteristics of a CYP116B1 substrate and the reaction it undergoes. Of interest were substrate binding in the active site, and stimulation of enzyme turnover and the products of any reaction. Further experiments were carried out on the isolated haem domain of CYP116B1 (HDCYP116B1) to probe the binding ofazole inhibitors, and to measure the redox potential of the haem and the effect of substrate upon these features. These experiments were carried out in order to characterise the P450 reaction centre and how these properties would likely affect the reaction with the substrate. In addition, the kinetics of the isolated reductase domain of CYP116B1 (RDCYP116B1) were investigated using artificial electron acceptors. This set of experiments was aimed at measuring the efficiency of 'electron throughput' of this domain of the enzyme for comparison with the overall P450 reaction kinetics.

5.1.3. Substrate Library

A library of potential substrates was assembled by looking at existing substrates both for CYP116B1 itself and for homologous enzymes identified in the work in section 4.1.2 (detailed above). Among the latter category were atrazine, as CYP116A1 has activity against this compound, and vanillic acid, as some vanillate-*O*-oxidases are supplied with electrons by PDOR-like reductases. As these enzymes are, potentially, evolutionarily related to CYP116B1 these substrates could help to probe this

relationship and the development of substrate specificity. The substrate library is listed in Table 5.1.

Substrate Library	
Trivial name	Systematic name
Substrates of Homologous enzymes	
3-chloroaniline	1-amino-3-chlorobenzene
vanillic acid	4-hydroxy-3-methoxybenzoic acid
4-chlorocatechol	4-chlorobenzene-1,2-diol
atrazine	1-chloro-3-ethylamino-5-isopropylamino-2,4,6-triazine
cartap	<i>S,S'</i> -[2-(dimethylamino)propane-1,3-diyl] bis(thiocarbamate)
Thiocarbamates	
Alkyl	
EPTC	<i>S</i> -ethyl dipropylthiocarbamate
vernolate	<i>S</i> -propyl dipropylthiocarbamate
pebulate	<i>S</i> -propyl butyl(ethyl)thiocarbamate
butylate	<i>S</i> -ethyl bis(2-methylpropyl)thiocarbamate
Cyclic Alkyl	
cycloate	<i>S</i> -ethyl cyclohexyl(ethyl)thiocarbamate
molinate	<i>S</i> -ethyl azepane-1-carbothioate
isopolinate	<i>S</i> -(1-methylethyl) azepane-1-carbothioate
Aromatic	
esprocarb	<i>S</i> -benzyl (1,2-dimethylpropyl)ethylthiocarbamate
methiobencarb	<i>S</i> -(4-methoxybenzyl) diethylthiocarbamate
prosulfocarb	<i>S</i> -benzyl dipropylthiocarbamate
fenothiocab	<i>S</i> -(4-phenoxybutyl) dimethylthiocarbamate
thiobencarb	<i>S</i> -(4-chlorophenyl) diethylthiocarbamate
orbencarb	<i>S</i> -(2-chlorobenzyl) diethylthiocarbamate
pyributicarb	<i>O</i> -(4- <i>tert</i> -butylphenyl)(6-metoxypyridin-2-yl)methylthiocarbamate
tiocarbazil	<i>S</i> -benzyl <i>bis</i> (1-methylpropyl)thiocarbamate
Chlorinated	
sulfallate	<i>S</i> -(2-chloroprop-2-en-1-yl)diethylthiocarbamate
di-allate	<i>S</i> -[(2 <i>Z</i>)-2,3-dichloroprop-2-en-1-yl] <i>bis</i> (2-methylpropyl)thiocarbamate
tri-allate	<i>S</i> -(2,3,3-trichloroprop-2-en-1-yl) <i>bis</i> (1-methylethyl)thiocarbamate
Miscellaneous	
prothiocarb	<i>S</i> -ethyl[3-(dimethylamino)propyl]thiocarbamate
methasulfocarb	4-[(methylcarbamoyl)sulfanyl]phenylmethanesulfonate

Table 5.1. Library of potential substrates for CYP116B1. The substrates were chosen based on two categories: substrates of homologous enzymes and thiocarbamates, as the two identified substrates for CYP116B1 are representative of this class of compound. The availability of substrates was checked with suppliers and available substrates were organised by functional similarities.

From the substrate library, a minimal panel of eight substrate candidates was chosen for the initial work. This panel was assembled to be representative of the various types of thiocarbamates in the substrate library and the substrates of homologous

enzymes. Chemical structures of the members of the panel are presented in Figure 5.2.

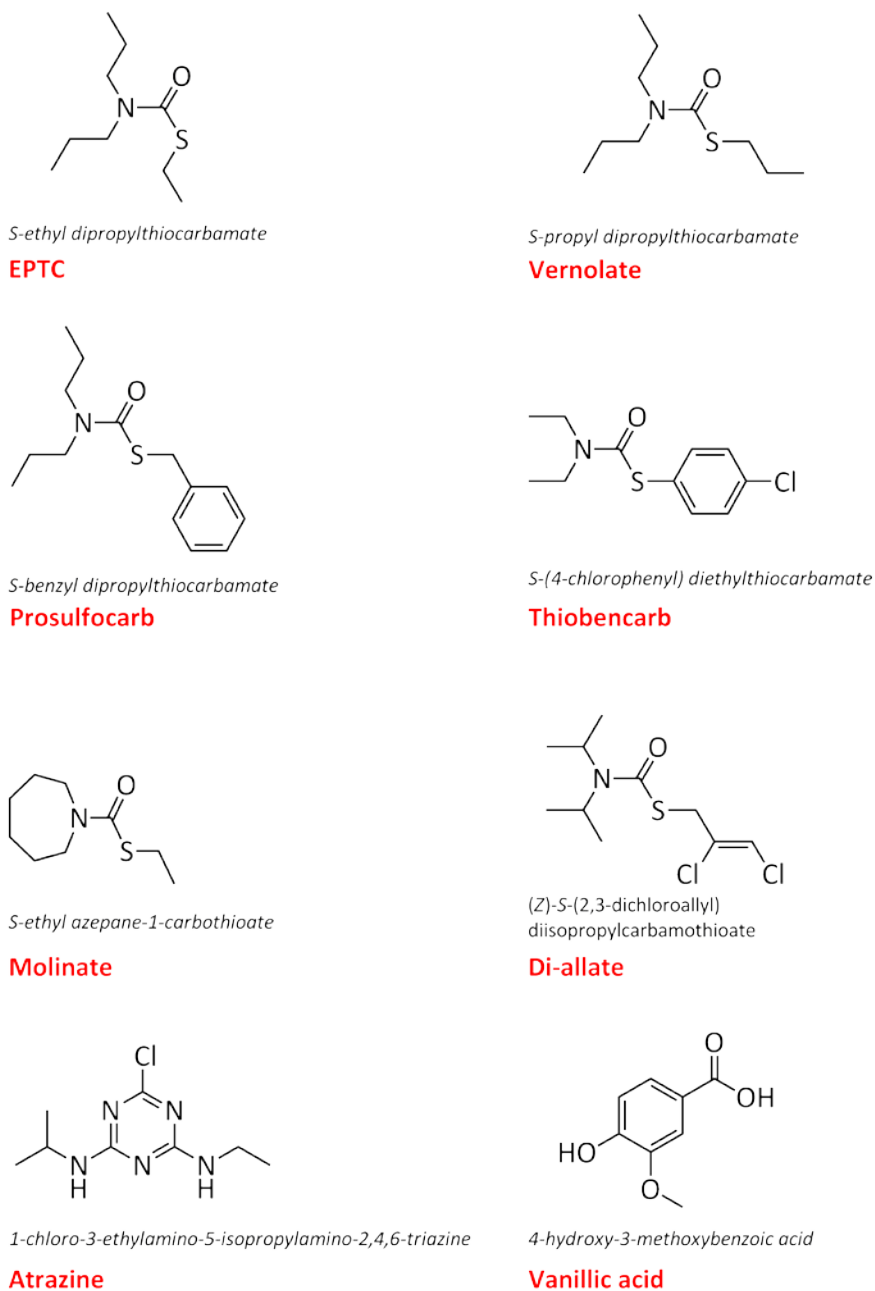


Figure 5.2. Structures of the eight CYP116B1 substrate candidates selected from the substrate library. The trivial names are given in red and systematic names in black.

5.2. Binding Studies

Substrate and inhibitor binding were investigated by performing binding titrations upon HDCYP116B1, as were carried out for HDCYP102A3 (see section 3.2.2). The titrations were performed as described in section 2.4.12.

5.2.1. Substrate Binding

Binding titrations were carried out using the two previously identified substrates for CYP116B1. The titrations were performed upon HDCYP116B1 ($\sim 5 \mu\text{M}$) using stocks of 100 mM EPTC and vernolate diluted in ethanol. Initially, additions of 0.1 μL of the substrate stocks were made, which equates to 10 μM step-increases of concentration in a 1 mL assay volume. When no change was observed, additions of 0.5-1 μL (~ 50 and 100 μM steps) were made. Even with these large additions relatively minor shifts in the Soret peak were observed and a binding curve could not be plotted. Spectra of HDCYP116B1 before and after addition of substrates EPTC and vernolate are shown in Figure 5.3. Both sets of spectra show a slight shift in the Soret peak towards a shorter wavelength (a type I shift). In addition, in both sets of spectra can be seen a small change in the relative intensities of the α/β peaks, which is likely indicative of the formation of high spin haem[205].

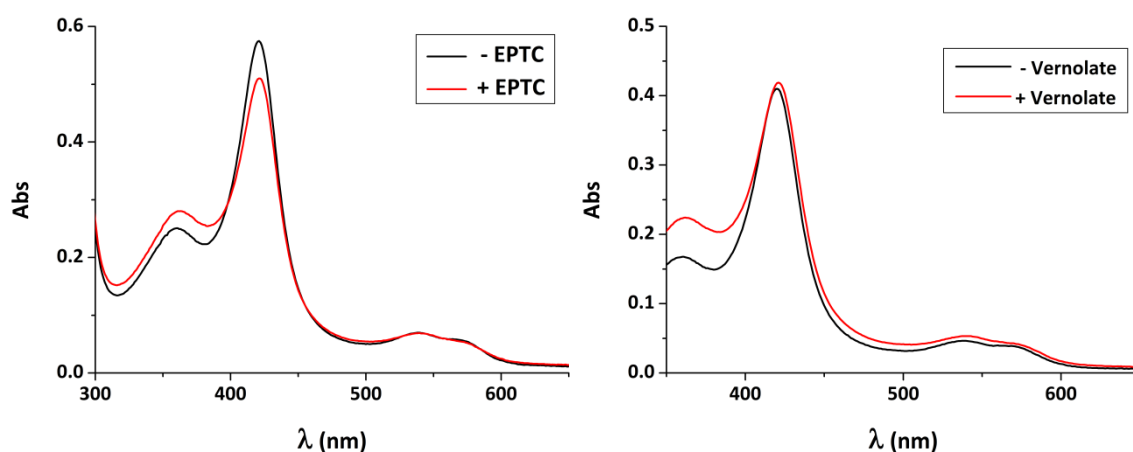


Figure 5.3. UV/visible spectra showing slight changes in haem Soret band of HDCYP116B1 (4 - 6 μM) in the presence of substrates EPTC (600 μM) and vernolate (500 μM). In both cases the changes are slight but there does seem to be a shift of the Soret peak to a shorter wavelength, indicating the production of a high spin haem species caused by the binding of substrate to the protein. Small changes are also notable in the α/β bands, and particularly for the EPTC titration.

5.2.2. Inhibitor Binding

Binding titrations of imidazole inhibitors were carried out on HDCYP116B1 as described previously (section 2.4.12) and elaborated in the previous chapter (section 3.2.2.2). The same set of inhibitors used in the work on HDCYP102A3 was used for CYP116B1 (imidazole and 1-, 2- and 4-phenylimidazoles) and as for this work difference spectra were constructed, wavelengths where the greatest change in absorbance occurred were identified and trough absorbance subtracted from peak absorbance. The resultant values were then plotted against ligand concentration to create a binding curve.

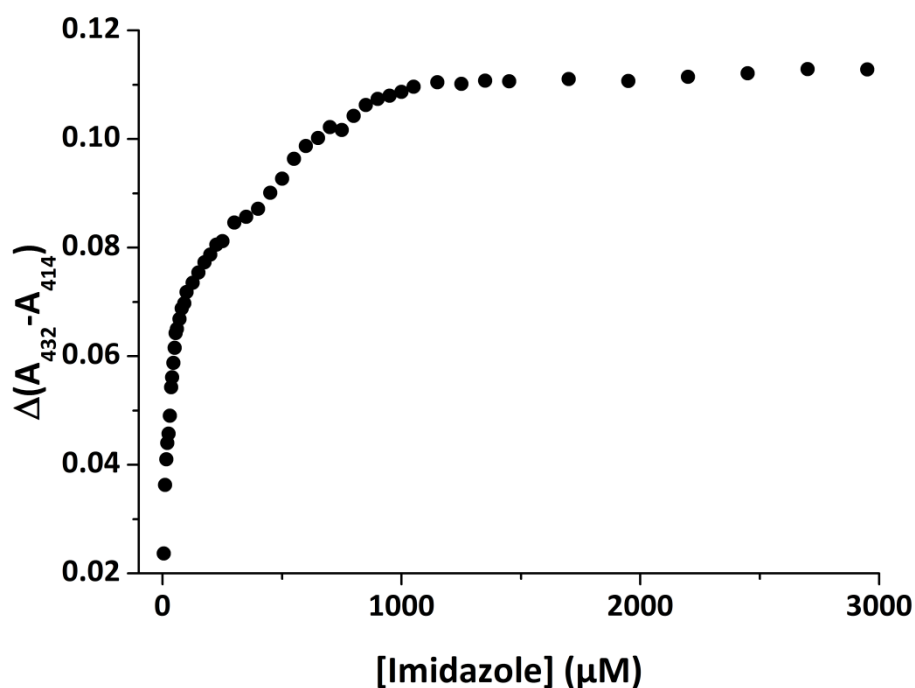


Figure 5.4. Binding curve for imidazole (5 – 1950 μM) titrated against HDCYP116B1 (5 μM). These data are biphasic and could not be fitted to any of the binding models detailed in section 1.2.2.

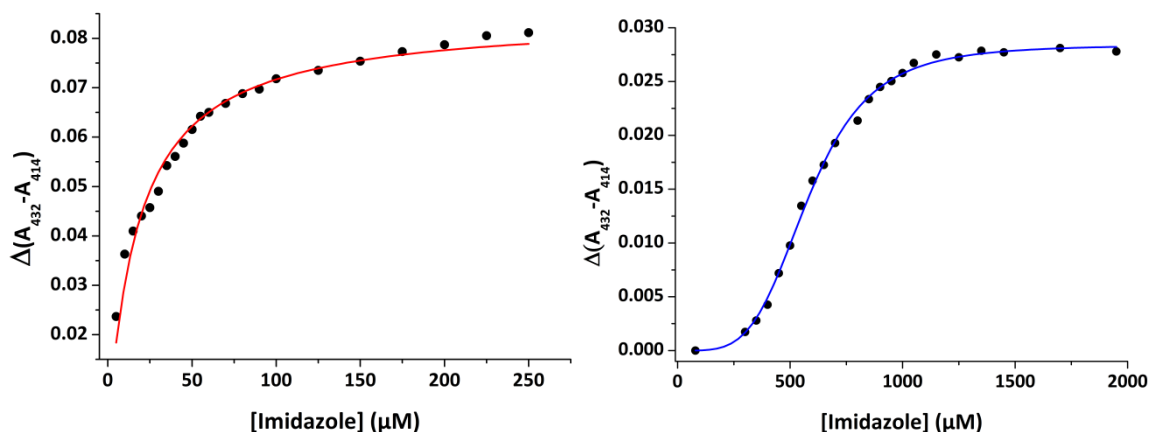


Figure 5.5. Binding data for imidazole titrated against HDCYP116B1 (5 μM) separated into two sections. The first (5 -250 μM imidazole) is fitted using the one-site hyperbolic binding equation (left). The second section (300 – 1950 μM imidazole) is fitted using the Hill function (right). K_d (K_H) values associated with these fits are presented in Table 5.2.

The binding curve for imidazole (Figure 5.4) could not be fitted directly to any one-site binding equation and seemed to show two distinct phases. These data were thus separated into two curves (Figure 5.5). The initial portion of the data (5 – 250 μM) could be fitted using a hyperbolic one-site binding equation and showed imidazole being bound tightly in an initial phase ($K_d = 18 \pm 1 \mu\text{M}$). The second part (300 – 1950 μM), appeared sigmoidal and could be fitted using a Hill function to derive an apparent dissociation constant (K_H). This portion of the data showed much weaker binding ($K_H = 580 \pm 5 \mu\text{M}$), perhaps indicating binding of two (or more) molecules of imidazole to HDCYP116B1. This tight imidazole binding provides some explanation for the residual imidazole-bound populations identified after purification of HDCYP116B1 and FLCYP116B1 by 6xHis tag affinity chromatography (see section 4.3.1). Though the question of why dilution and dialysis were unable to remove this bound imidazole still remains.

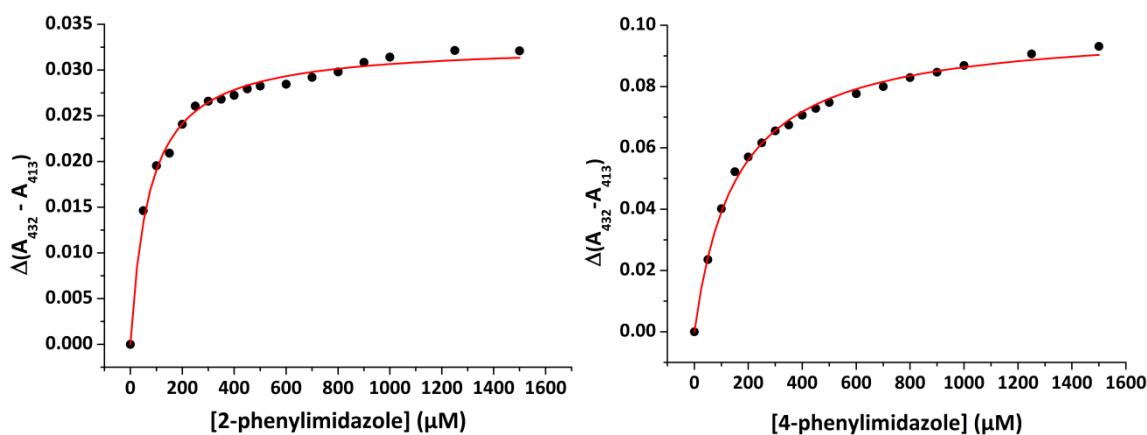


Figure 5.6. Binding curves for HDCYP116B1 ($\sim 5 \mu\text{M}$) with (left) 2- and (right) 4-phenylimidazoles. Both were fitted using a hyperbolic one-site binding equation (section 1.2.2, Equation 1.4). K_d values associated with these fits are presented in Table 5.2.

The binding curves for 2- and 4-phenylimidazoles (Figure 5.6) could both be fitted to a hyperbolic one-site binding equation and produced K_d values of $72 \pm 5 \mu\text{M}$ and $155 \pm 7 \mu\text{M}$ respectively. This shows that 2-phenylimidazole is slightly more tightly bound to HDCYP116B1 than to HDCYP102A3 ($107 \pm 23 \mu\text{M}$, see section 3.2.2.2), and than to CYP51B1 and CYP121A1 from *M. tuberculosis* ($K_d = 3190 \pm 370 \mu\text{M}$ [206] and $101.8 \mu\text{M}$ [207] respectively) among others[208]. HDCYP116B1 may represent one of the tightest 2-phenylimidazole binding P450 enzymes yet seen. Many of the examples above of large K_d values for 2-phenylimidazole are attributable to a binding mode where the ligand does not directly ligate the haem. This is evidenced by the structure of CYP101A1 (P450cam) with this ligand bound[157]. In these examples the ligand does not cause a spectral shift upon binding, due to restricted access of the ligating nitrogen. In the case of HDCYP116B1 a type II shift is seen ($413 \rightarrow 432 \text{ nm}$ in the difference spectra), possibly suggesting, that the substrate binding cavity of this enzyme can accommodate 2-phenylimidazole in such a position as to allow direct ligation of the haem.

K_d values for all inhibitors, along with associated errors derived from the fitting, are presented in Table 5.2.


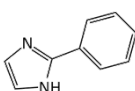
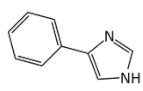
Ligand		K_d/K_H^* (μM)	
Imidazole		18 \pm 1	580 \pm 5*
2-phenylimidazole			72 \pm 5
4-phenylimidazole			155 \pm 7

Table 5.2. Dissociation constants (K_d/K_H) for a range of imidazole inhibitors binding to HDCYP116B1. The K_H value for imidazole is indicated with an asterisk.

5.3. Redox Potentiometry

Potentiometric studies were performed for FLCYP116B1 and HDCYP116B1 to investigate the influence of substrate binding and the presence of the reductase domain upon the haem iron redox potential of the enzyme. Additionally, by comparison with homologous proteins (*i.e.* other P450 enzymes and PDOR-type enzymes) the redox potentials can be correlated and the influence of the redox system established by the fusion of these two proteins upon the bound cofactors can be examined.

Attempts to determine a redox potential for FLCYP116B1 were stymied by aggregation of the protein during the experiment. The cause of this aggregation is unknown as experimental conditions such as dithionite reduction had not previously caused aggregation in other experiments. The mediator compounds added may have had an adverse interaction with the FLCYP116B1 protein, as too might the requirement for extended incubation of the enzyme at 25 °C for the duration of the redox titration. As a result no redox potential measurements could be recorded for the full length enzyme.

The HDCYP116B1 protein was amenable to potentiometric studies and experiments were carried out as described in section 2.4.13. The haem iron $\text{Fe}^{2+}/\text{Fe}^{3+}$ midpoint potential was determined for HDCYP116B1 in the presence and absence of substrate EPTC (1 mM). As with HDCYP102A3, the UV/Visible spectrum was recorded after small additions of sodium dithionite solution while the redox potential against a calomel electrode (+244 mV vs. SHE) was measured. For both the substrate free and

substrate bound samples, absorbance measurements at 559 nm (the α -band) were plotted against the standard electrode potential at which they were recorded. The resulting potential curves (Figure 5.7) were fitted using the Nernst equation for one electron transfer.

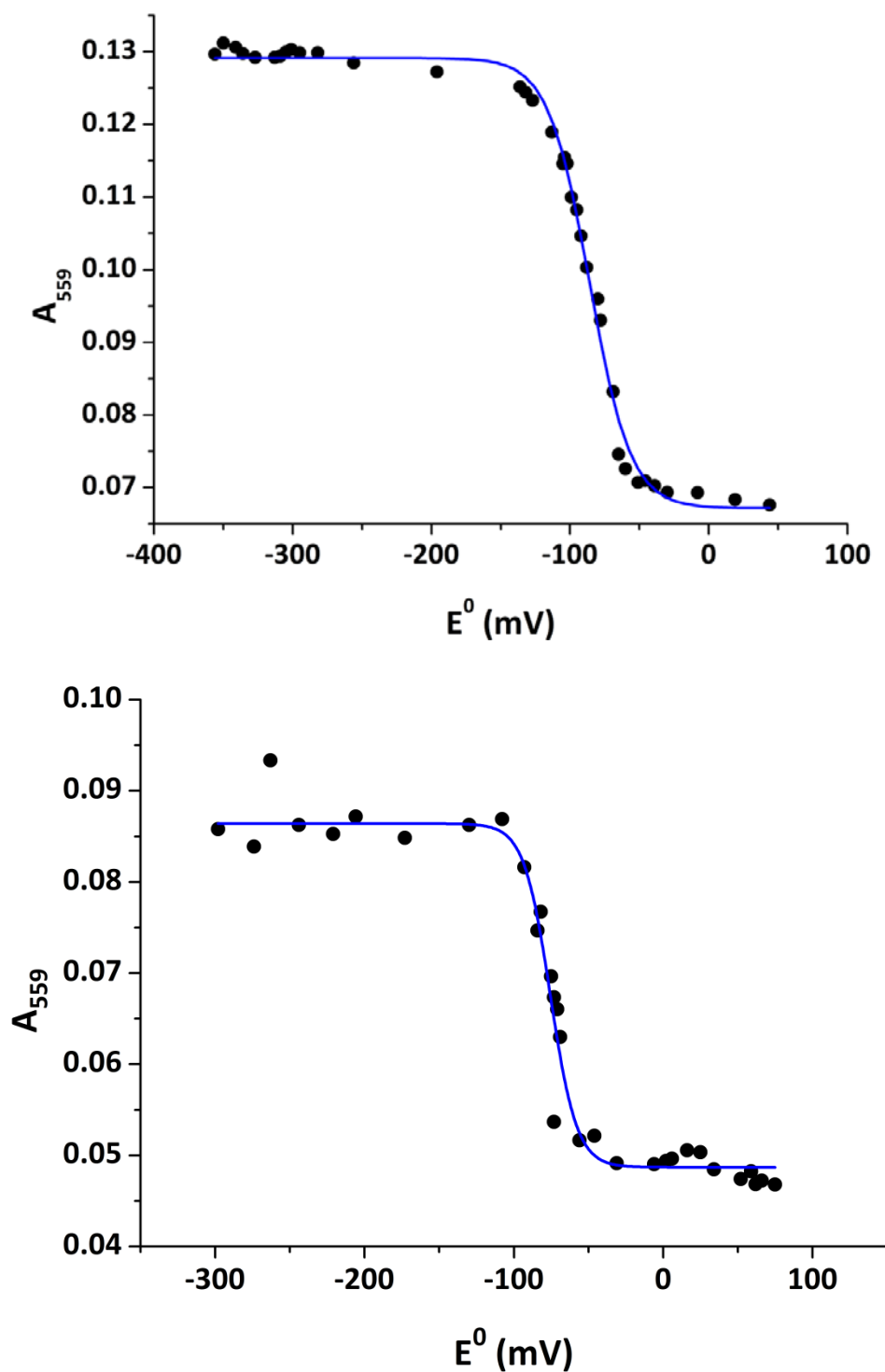


Figure 5.7. Redox potential curves for the haem domain of CYP116B1 as purified (*i.e.* low-spin, top) and in the presence of substrate EPTC (bottom). The x-axis is the potential against the standard hydrogen electrode and the y-axis is the change in haem absorbance at 559 nm between the α/β bands. Both data sets are fitted using the Nernst equation for one electron transfer (blue lines) to determine midpoint potentials: -85 ± 1 mV (substrate free) and -75 ± 2 mV (EPTC bound). The experiments were carried out as described previously (section 2.4.13)

The midpoint potentials ($\text{Fe}^{2+}/\text{Fe}^{3+}$) of the haem iron in the substrate-free and substrate-bound forms of the HDCYP116B1 protein were -85 ± 1 mV and -75 ± 2 mV respectively. This represents a maximum shift of $\sim +13$ mV (within error limits) caused by the presence of substrate. The magnitude of this shift is around ten times smaller than substrate induced shifts recorded for other P450s including P450cam (~ 130 mV[164]), BM3 (~ 100 mV[111]), and certainly that recorded for HDCYP102A3 in the presence of NPG (~ 210 mV, see section 3.2.5). This is consistent with the rather small spectral change induced by substrate binding (EPTC, vernolate) upon the HDCYP116B1 Soret peak. The protein is predominantly low-spin ferric as purified, and there are only very minor changes induced by substrate binding, likely reflecting a shift towards high-spin ferric of only $\sim 5\%$ at most (section 5.2.1).

The most notable point is how positive the haem iron potential is in HDCYP116B1 (~ 200 mV more positive than the majority of other P450s analysed). At this stage it is not possible to ascertain whether this is a consequence of the removal of the haem domain from the intact CYP116B1 enzyme, but clearly further work is required to establish whether the haem iron potential is substantially altered in the intact CYP116B1 enzyme and to compare the haem potential with those of the FMN and [2Fe-2S] centres in the enzyme.

5.4. Steady State Kinetics of FLCYP116B1

Steady-state kinetic studies were carried out for each of the substrate candidates to determine if there is a dependency of NADPH oxidation rate upon their concentration using CYP116B1. Stimulation of NADPH oxidation could indicate substrate-like binding and/or the oxidation of these molecules by the P450. Assays were carried out as described (section 2.4.16) with NADPH oxidation measured by a decrease in absorbance at 340 nm ($\Delta\epsilon_{340} = 6.21 \text{ mM}^{-1} \text{ cm}^{-1}$). A range of concentrations of each substrate candidate was tested (typically in the range 0.1 – 3.0 mM) and each assay was run in triplicate, and from these an average value of the rate was taken. Rates were corrected for enzyme concentration and then plotted against concentration of substrate. The plotted data were then fitted using either the Michaelis-Menten function or the Hill function in cases where an apparent sigmoidal dependence of NADPH oxidation on substrate concentration was observed.

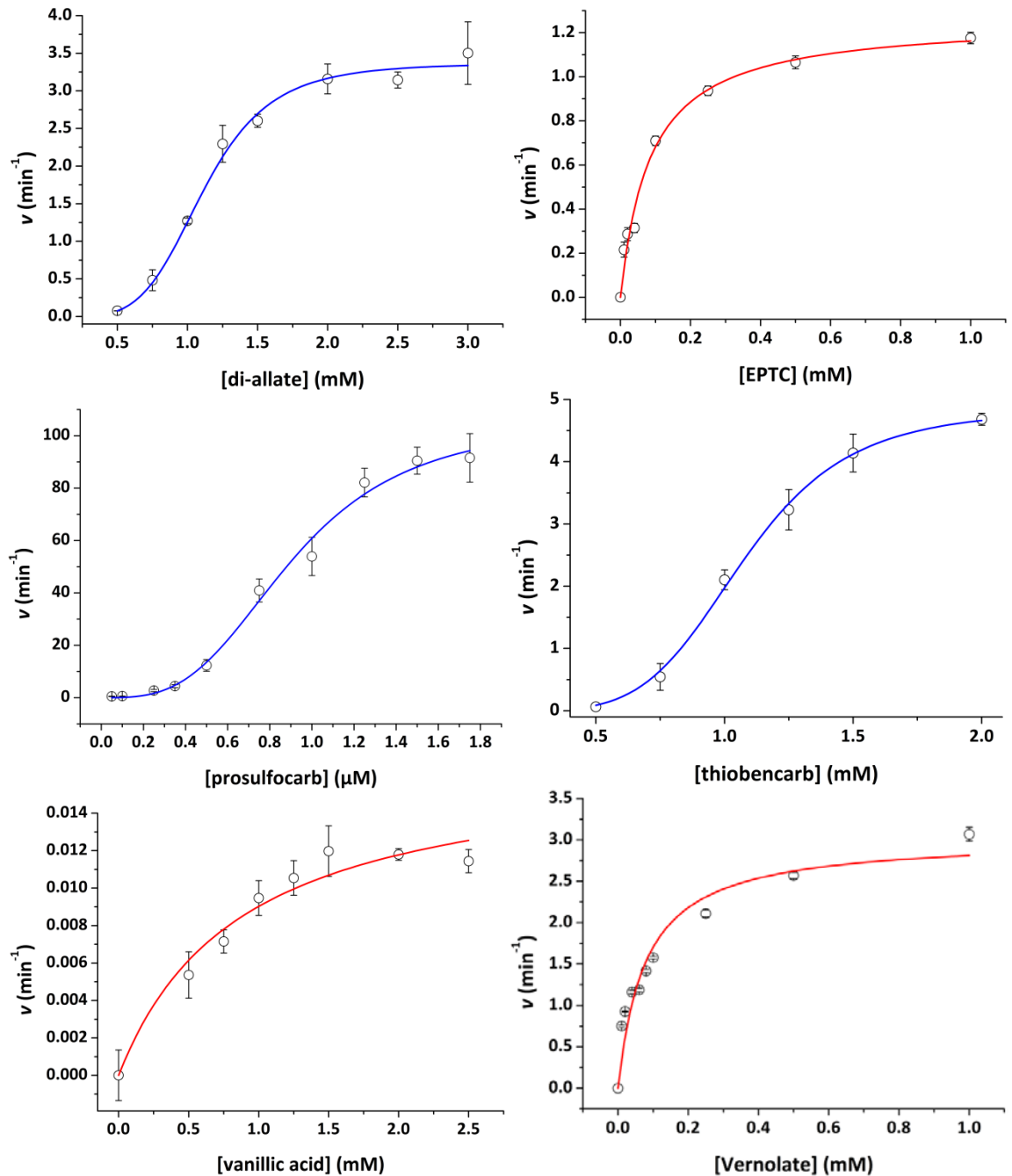


Figure 5.8. Kinetic curves for the dependency of FLCYP116B1-dependent NADPH oxidation on the substrate concentration, using various substrate candidates for the enzyme. Each data point is the mean average of values recorded for three assays and the error bars represent the standard deviation for these values. Data are fitted using either the Michaelis-Menten (red line) or Hill (blue line) equations in cases of data showing a hyperbolic or sigmoidal dependency, respectively. Hill numbers (n) associated with the fitting are 4.7 ± 0.6 (di-allate), 3.2 ± 0.5 (prosulfocarb) and 5.2 ± 0.5 (thiobencarb).

For six of the substrate candidates, a concentration dependent stimulation of the rate of NADPH oxidation by FLCYP116B1 was detected suggesting that they are indeed bound to the enzyme in the manner of substrates. Kinetic curves for these six molecules (di-allate, EPTC, prosulfocarb, thiobencarb, vanillic acid and vernolate) are shown in Figure 5.8 and the associated kinetic constants are presented in Table 5.3. For the remaining two substrate candidates (atrazine and molinate) no stimulation of the rate of NADPH oxidation was detected, suggesting that neither of these is a substrate for CYP116B1 and that the rates of NADPH oxidation observed in their presence are simply those for reduction of molecular oxygen through flavin and [2Fe-2S] centres in the enzyme.

Atrazine is the least similar to the other substrate candidates as it does not contain a carbonyl functional group. Homologous enzymes, that are likely to be evolutionarily related to CYP116B1, are capable of reaction with atrazine. It appears, therefore, that this capacity is not retained in CYP116B1.

Molinate is a thiocarbamate like most of the other substrate candidates, although it also has a large, saturated, seven-membered ring which is unique among them. It is possible that this group hinders binding of molinate to CYP116B1, and/or turnover of this molecule.

For three of the six substrate candidates stimulating NADPH oxidation (di-allate, prosulfocarb and thiobencarb) a sigmoidal dependency of the rate of NADPH consumption was observed. From the fitting of kinetic curves for these substrates using the Hill equation, the Hill numbers are 4.7 ± 0.6 , 3.2 ± 0.5 and 5.2 ± 0.5 respectively for di-allate, prosulfocarb and thiobencarb. This describes an apparent cooperativity in the stimulation of NADPH oxidation for these substrate candidates. The reasons are likely similar to those given for cooperativity in the binding of substrates to HDCYP102A3 (see section 3.2.2.1). In short, the binding of more than one molecule of the substrate candidate in two distinct sites within the enzyme, or alternatively the substrate induced oligomerisation of enzyme monomers. In each case, the apparent cooperativity shown in concentration dependency experiments is supposed to be produced by one of these phenomena.

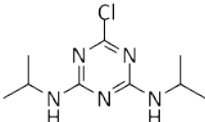
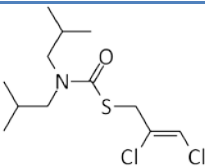
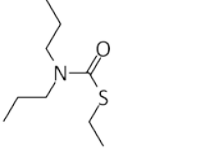
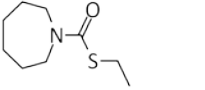
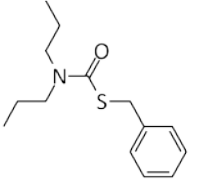
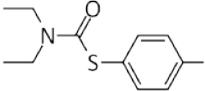
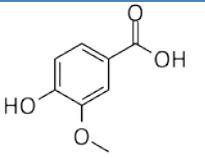
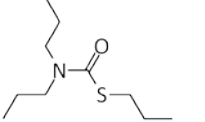
Substrate Candidate	k_{cat} (min^{-1})	K_{M} (μM)
Atrazine 	-	-
Di-allate* 	3.4 ±0.1	1120 ±30
EPTC 	1.3 ±0.1	85 ±13
Molinate 	-	-
Prosulfocarb* 	106 ±8	0.90 ±0.06
Thiobencarb* 	4.8 ±0.2	1070 ±20
Vanillic acid 	0.017 ±0.002	900 ±300
Vernolate 	3.0 ±0.2	78 ±17

Table 5.3. Kinetic constants (k_{cat} and K_{M}) for FLCYP116B1 based on NADPH consumption in the presence of a range of substrates. The values of these constants are derived from fitting of kinetic curves using the Michaelis-Menten equation or the Hill equation (indicated by *). Error values quoted are based on the errors from the fitting process. Chemical structures of the substrates are shown next to the names.

The substrate candidate with the highest rate stimulation is prosulfocarb ($k_{\text{cat}} = 106 \pm 8 \text{ min}^{-1}$), which is a previously unknown substrate for CYP116B1. The low value of the K_{M} ($0.90 \pm 0.06 \mu\text{M}$) suggests that there is strong binding affinity for this substrate of the enzyme. This may be due to the benzyl group present in prosulfocarb which could have a strong aromatic stacking interaction with Phe-86 which, by alignment with other P450s, would be predicted to lie in the CYP116B1 substrate binding cavity over the haem.

Di-allate contains a fairly polar substituent group (1,2-*cis*-dichloropropenyl) attached to the sulfur of the thiocarbamate group, which may explain the high K_M value ($1120 \pm 30 \mu\text{M}$) for this substrate. Although increasing polarity of the molecule, this group is also highly reactive and this may explain the rate of reaction achieved ($k_{\text{cat}} = 3.4 \pm 0.1 \text{ min}^{-1}$), comparable with established substrate vernolate ($k_{\text{cat}} = 3.0 \pm 0.2 \text{ min}^{-1}$). The increased polarity is also mitigated in part by the *N*-substituent isobutyl groups.

The 4-chlorophenyl group of thiobencarb gives it a combination of the properties of prosulfocarb and di-allate – an aromatic group conducive to substrate binding to the enzyme and a chloro substituent increasing overall polarity. The *N*-substituent ethyl chains are also reduced in length compared to prosulfocarb (*n*-propyl) and di-allate (isobutyl). Nevertheless, the rate achieved ($k_{\text{cat}} = 4.8 \pm 0.2 \text{ min}^{-1}$) is also comparable to that for vernolate, perhaps highlighting the influence of the aromatic stacking interaction mentioned above for prosulfocarb.

5.5. Product Determination

These experiments sought to determine the nature of the products of the CYP116B1 enzyme reaction with each of the substrate candidates. Overnight reactions were set up based on the method of Warman[135] containing the CYP116B1 enzyme, one of the substrate candidates, NADPH and an NADPH regeneration system (see section 2.4.20.1). For the substrates other than those previously established (EPTC and vernolate), control reactions were also carried out in which all components were added except the CYP116B1 enzyme. The products of the reaction were then purified, using solid-phase extraction columns (see section 2.4.20.2).

The products of the enzymatic reaction of CYP116B1 against each of the substrate candidates was examined by high performance liquid chromatography linked mass spectrometry (HPLC-MS) in an attempt to determine their nature. HPLC-MS experiments were carried out as described in section 2.4.20.3, with the assistance of Dr. Bettina Nestl of the Department of Chemistry, University of Manchester.

From these experiments two attributes could be determined: firstly the retention time of the product compound(s) on the HPLC column and secondly their mass spectrum profile. Elution from the HPLC column was monitored over time for absorbance at 254 nm & 220 nm, and for signal on the mass spectrum detector (MSD). The selected wavelengths are the default setting on the instrument, as a range of organic chromophores (such as aromatics and other conjugated groups) are likely to show absorbance at one or both of these wavelengths.

The resulting HPLC traces and mass spectra corresponding to elution peaks are presented in the sub sections below for the reaction of each substrate candidate with CYP116B1. In all HPLC traces an elution peak is visible at ~64 minutes, this is associated with the solvent (in which the extracted products are dissolved) eluting from the HPLC column and is not relevant to the product compounds themselves.

5.5.1. Atrazine

The HPLC traces for the products of the reaction of atrazine with CYP116B1 (Figure 5.9.) show elution peaks at ~1.2 and ~1.5 minutes. These peaks are visible in the both absorbance traces, and are also prominent in the MSD trace.

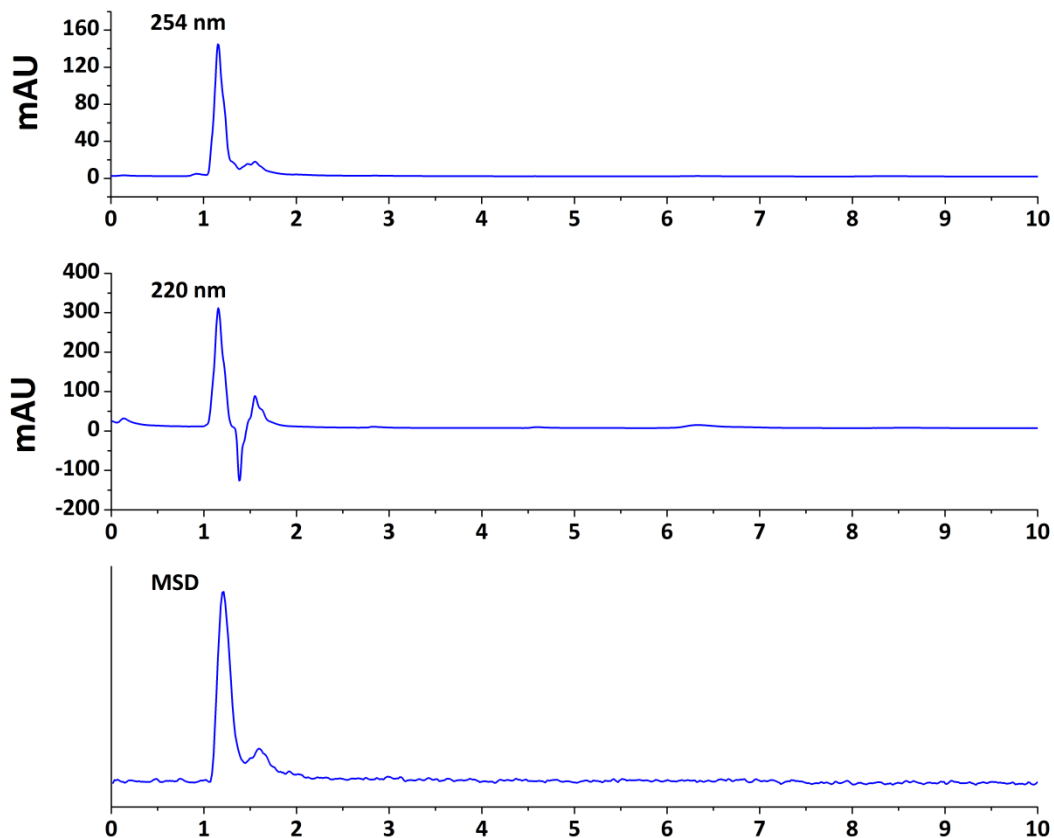


Figure 5.9. HPLC traces for the products of the reaction of atrazine with CYP116B1. Elution was monitored for absorbance at wavelengths of 254 and 220 nm and for signal on the mass spectrometric detector (MSD). Elution peaks occur in both the 254 nm and 220 nm absorbance traces and the MSD trace at ~1.2 and ~1.6 minutes.

The mass spectrum for the peak in the trace at 1.6 minutes (Figure 5.10) shows a peak at m/z 226.3. This corresponds exactly to a peak in the equivalent mass spectrum for the control reaction and is within 4 units of the mass of atrazine ($229.71 \text{ g mol}^{-1}$).

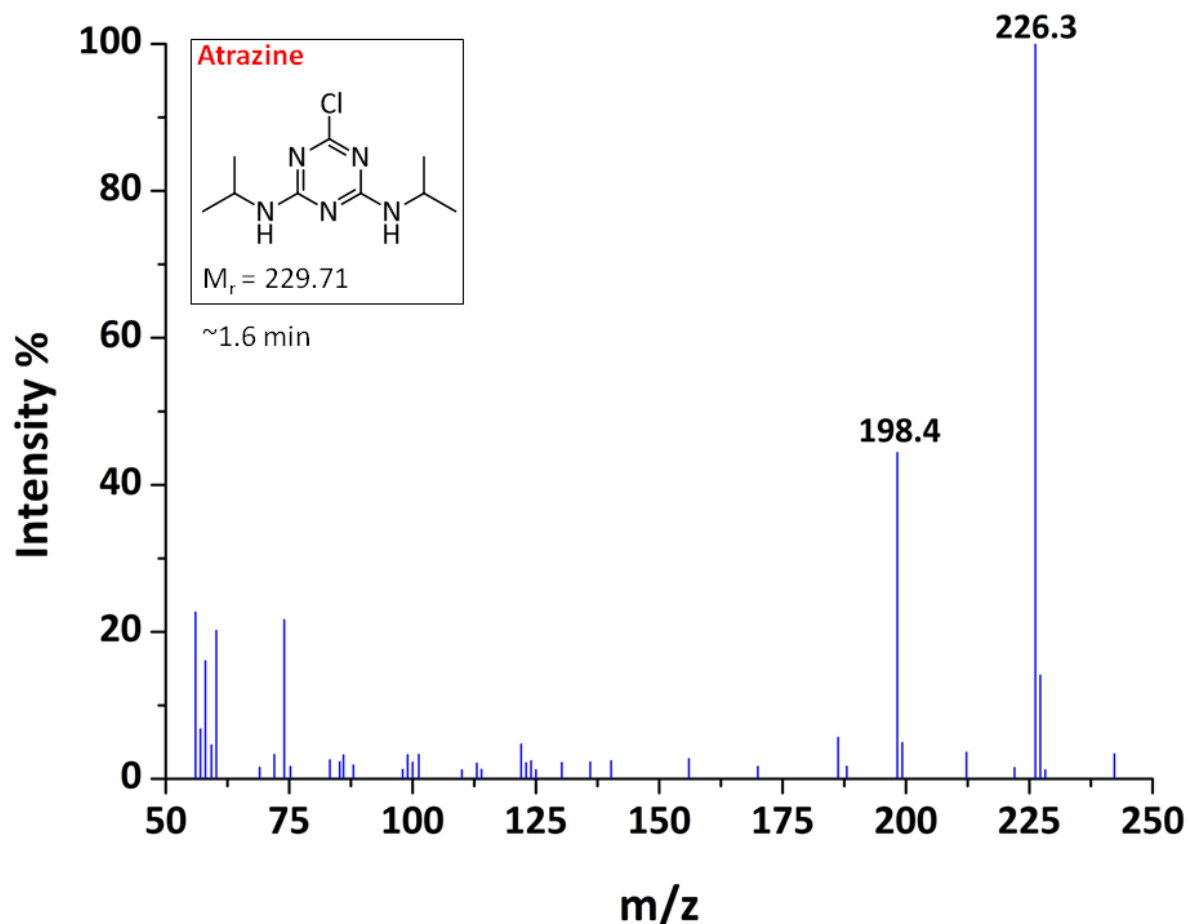


Figure 5.10. Mass spectrum of the purified products of a CYP116B1 turnover reaction in the presence of substrate candidate atrazine. The mass spectrum was recorded for a fraction of the product sample with a HPLC retention time of ~ 1.6 minutes. The spectrum shows a peak at a m/z ratio of 226.3, which is within 4 mass units of the molecular mass of atrazine ($229.71 \text{ g mol}^{-1}$).

Upon comparison, the mass spectrum in Figure 5.10 and that of the peak at ~ 1.2 min are largely identical to the corresponding mass spectra for the control reaction with no enzyme. This result suggests that there is no activity against atrazine by CYP116B1. This is supported by the observation from steady-state experiments (see section 5.4), of no stimulation of the rate of NADPH oxidation by atrazine.

5.5.2. Di-allate

The HPLC traces for the products of the reaction of di-allate with CYP116B1 (Figure 5.11) show prominent elution peaks at ~1.2 and ~8.8 minutes. These peaks are visible in the both absorbance traces, and the two earlier peaks are also visible in the MSD trace.

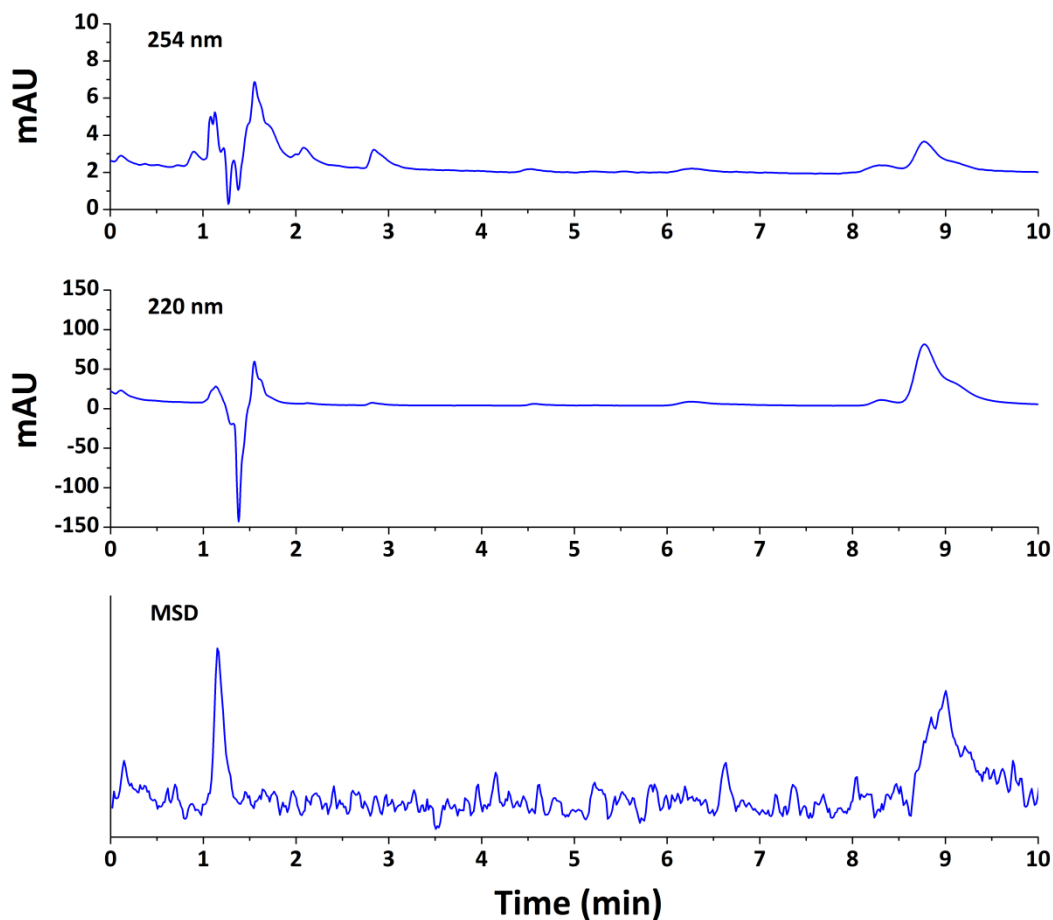


Figure 5.11. HPLC traces for the products of the reaction of di-allate with CYP116B1. Elution was monitored for absorbance at wavelengths of 254 and 220 nm and for signal on the mass spectrometric detector (MSD). Elution peaks occur in both the 254 nm and 220 nm traces at ~1.2 and ~8.8 minutes, with the first two peaks also visible in the MSD trace.

The mass spectrum for the peak in the trace at ~1.2 minutes is almost identical to the corresponding mass spectrum for the control reaction (Figure 5.12). This suggests either than no reaction has taken place, or that the products are not contained in this elution peak.

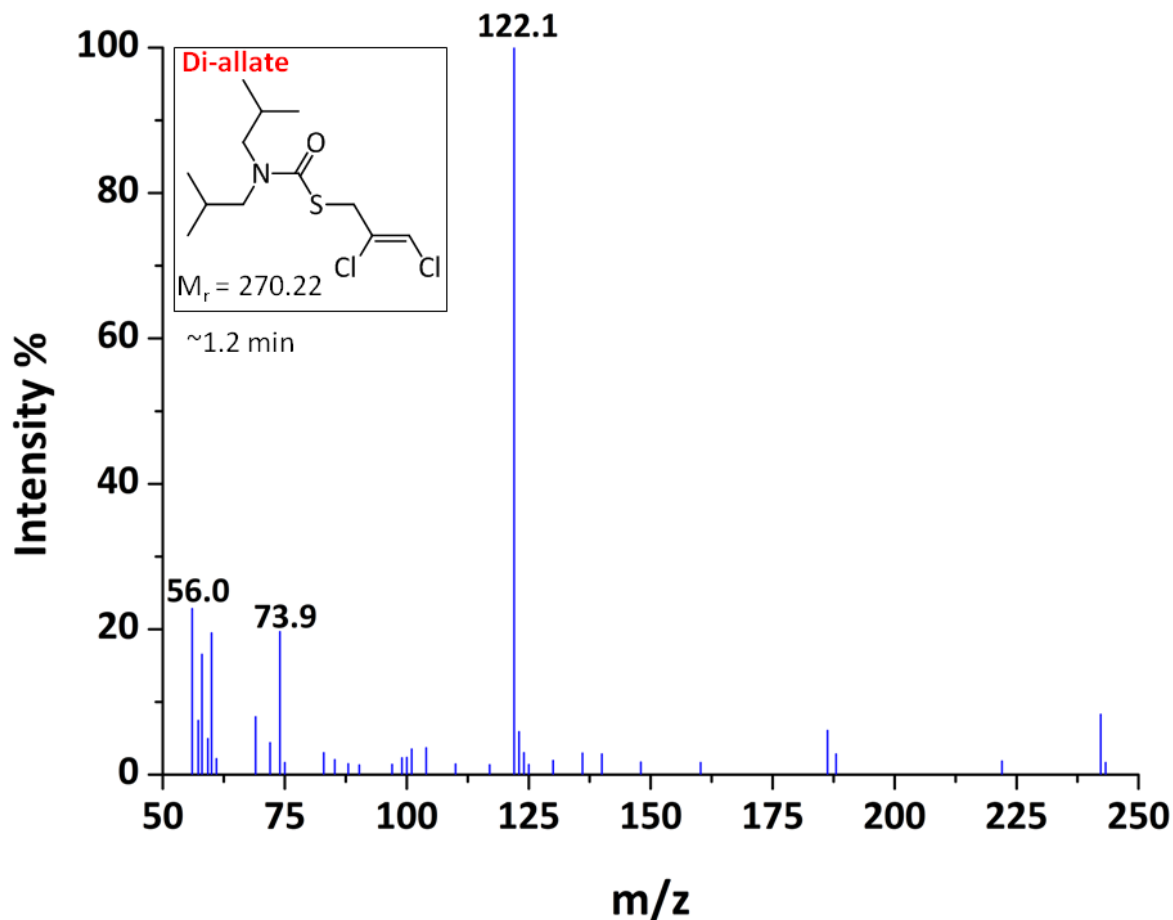


Figure 5.12. Mass spectrum of the purified products of a CYP116B1 turnover reaction in the presence of substrate candidate di-allate. The mass spectrum was recorded for a fraction of the product sample with a HPLC retention time of ~ 1.2 minutes. The equivalent mass spectrum for the control reaction is shown as an inset.

The mass spectrum of the elution at ~ 8.8 minutes shows a peak at m/z 270.0 which corresponds closely to the molecular weight of di-allate ($270.22 \text{ g mol}^{-1}$). The spectrum is also nearly identical to the corresponding spectrum for the control reaction. This indicates that the substrate is still present in the reaction product, further suggesting that no reaction has occurred.

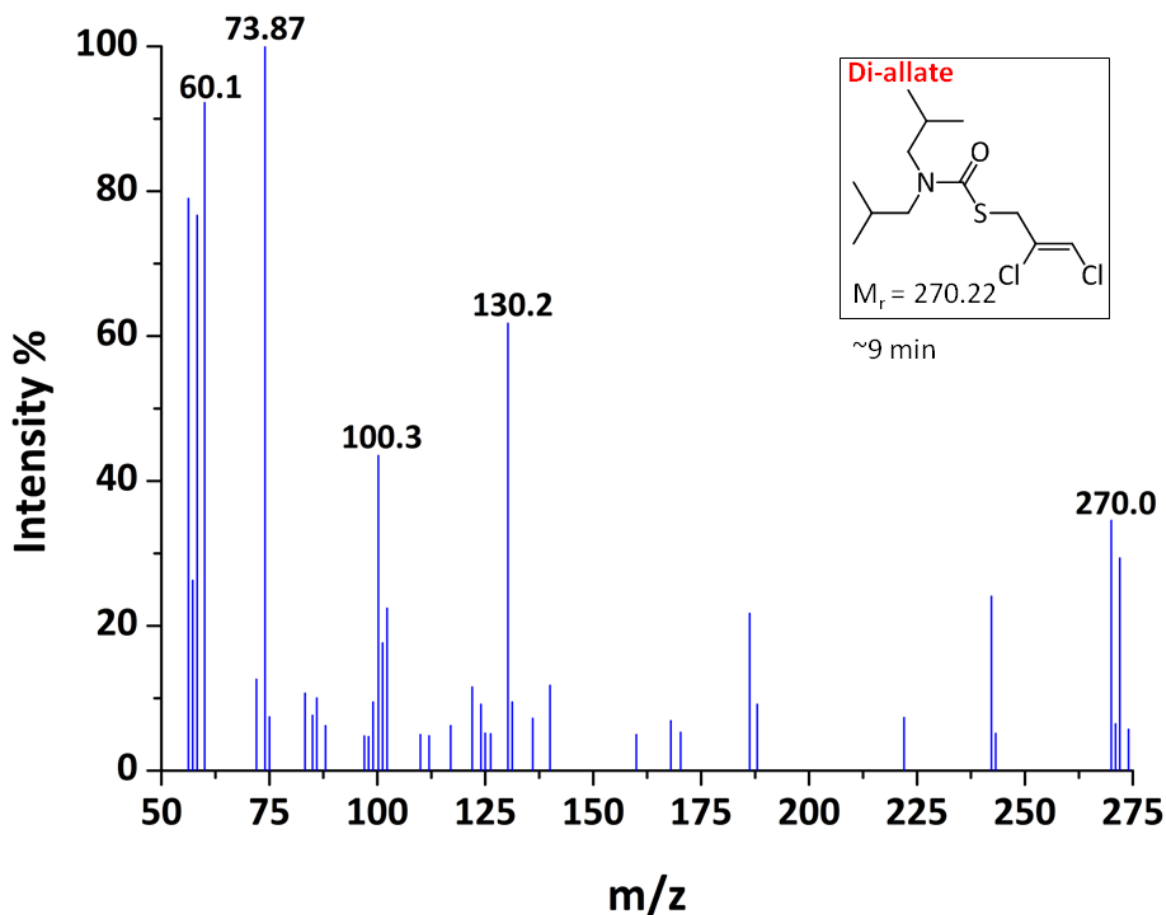


Figure 5.13 Mass spectrum of the purified products of a CYP116B1 turnover reaction in the presence of substrate candidate di-allate. The mass spectrum was recorded for a fraction of the product sample with a HPLC retention time of ~ 8.8 minutes. The equivalent mass spectrum for the control reaction is shown as an inset.

In the mass spectra for both HPLC elution peaks, many peaks are present at a range of m/z ratios. Di-allate does possess many reactive functional groups and its decomposition following S-oxidation has been investigated[209]. However, no potential products from mechanisms described in this work, or those from other mechanisms such as *N*-dealkylation, have molecular masses corresponding to any peaks in the mass spectra above. This may indicate that further breakdown of the reaction products has occurred after the initial reaction, either due to further P450 activity or spontaneous decomposition. Alternatively, the range of peaks in the mass spectrum could indicate the spontaneous breakdown of this substrate candidate molecule. This breakdown could potentially be caused by the aqueous environment of the enzyme reaction mixture in which the di-allate was present for 16 hours.

5.5.3. EPTC

The HPLC traces for the products of the reaction of EPTC with CYP116B1 (Figure 5.14) show elution peaks at ~3 and ~5 minutes. The peak at ~3 minutes is most intense in the 254 nm trace, and that at ~5 minutes in the 220 nm trace. The latter is also clearly visible while the former is ill-defined, in the MSD trace.

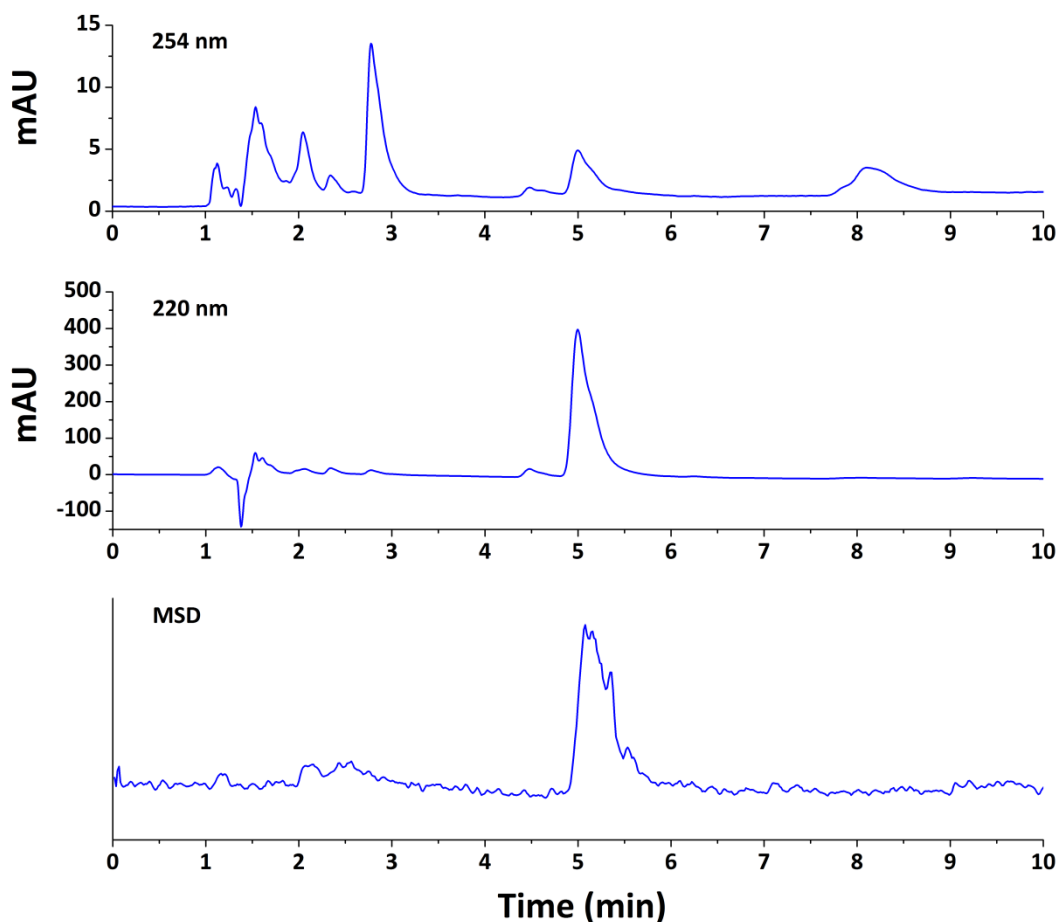


Figure 5.14. HPLC traces for the products of the reaction of EPTC with CYP116B1. Elution was monitored for absorbance at wavelengths of 254 and 220 nm and for signal on the mass spectrometer detector (MSD). The largest peaks occur in the 254 nm trace at 2.780 minutes and in the 220 nm trace at 4.997 minutes.

The mass spectrum for the peak at ~ 2.2 minutes shows a large peak at m/z 206.3 which corresponds to hydroxy EPTC, a product that has previously been observed for the reaction of CYP116B1 with this substrate. Further peaks at m/z ratios of 144.0 and 58.0 correspond fairly closely to *N*-depropyl EPTC ($M_r = 147.24 \text{ g mol}^{-1}$), and propanal ($M_r = 58.08 \text{ g mol}^{-1}$). These are the two proposed products of the decomposition of hydroxy EPTC by the species of *Rhodococcus* possessing the CYP116 enzymes and able to mineralise EPTC[202] (see Figure 5.1). This suggests that the decomposition of hydroxy EPTC, into these products, has occurred under the conditions of the reaction. However, as stated in section 5.5.8 below, the mass spectrum for the products of the CYP116B1 reaction with vernolate cast doubt upon this conclusion.

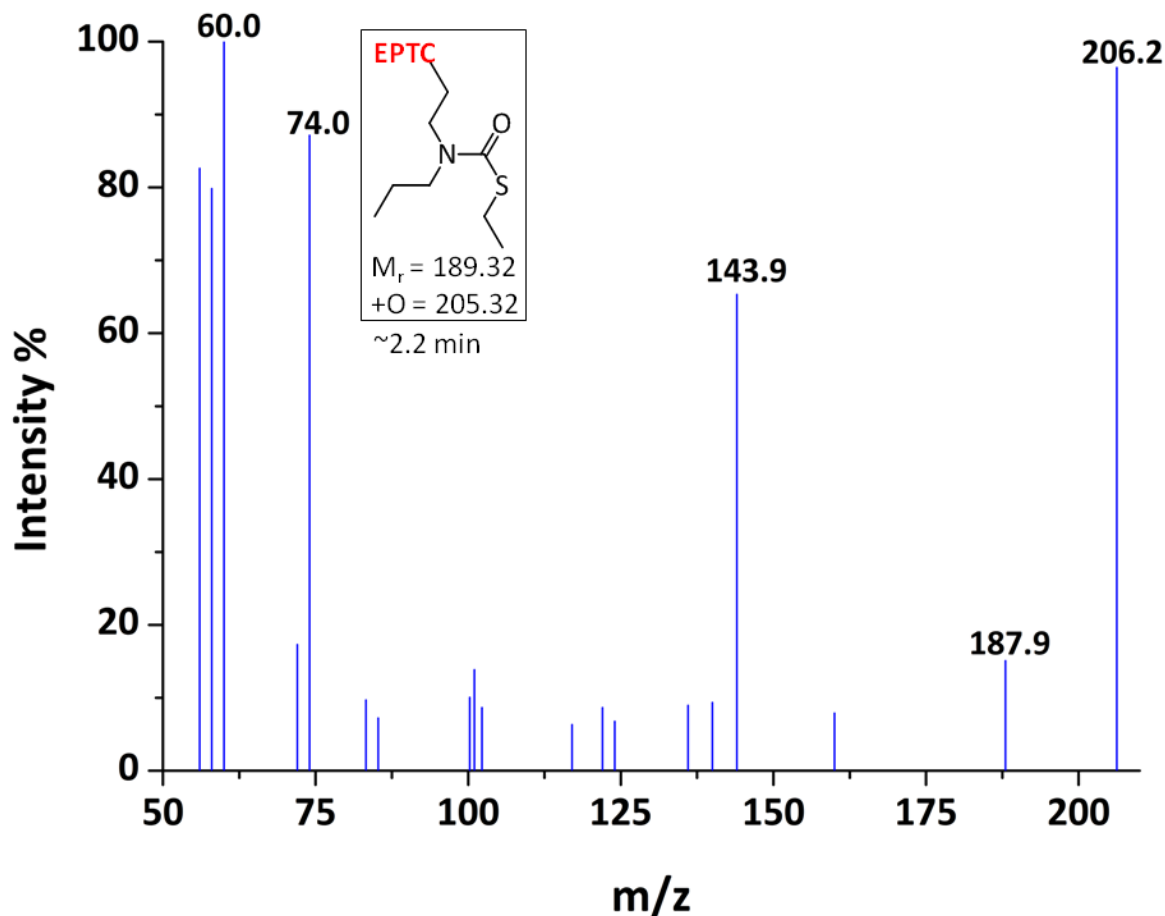


Figure 5.15. Mass spectrum of a CYP116B1 turnover reaction in the presence of substrate EPTC. The mass spectrum was recorded for a fraction of the reaction with a HPLC retention time of ~ 2.2 minutes. The spectrum shows a peak at a m/z of 206.3 which corresponds (within 1 mass unit) to a hydroxylated product. Additionally, peaks at m/z 144.0 and 58.0 correspond (within 3 mass units) to *N*-depropyl EPTC and propanal.

5.5.4. Molinate

The HPLC traces for the products of the reaction of molinate with CYP116B1 (Figure 5.16) show prominent elution peaks at ~1, ~4 minutes. Both elution peaks are visible in the 254 nm absorbance trace, and the latter in the 220 nm trace. Only the latter elution peak is present in the MSD trace.

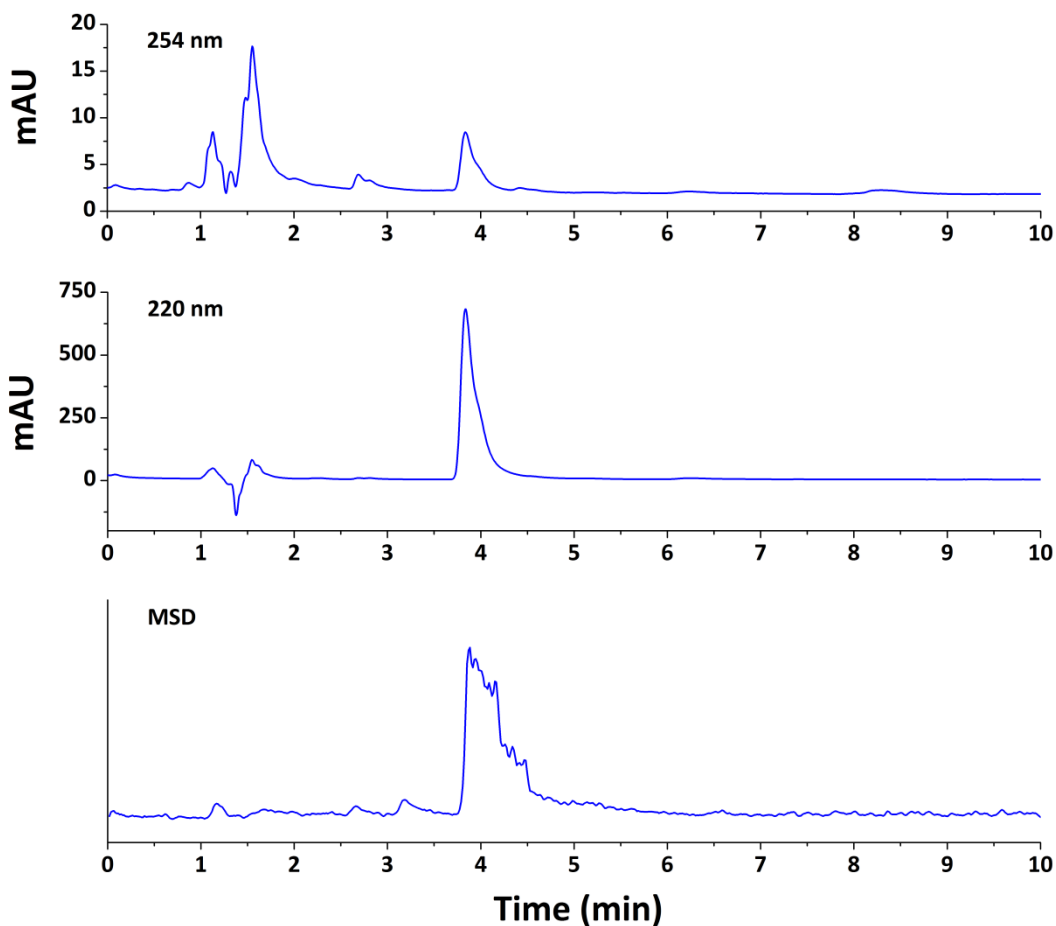


Figure 5.16 HPLC traces for the products of the reaction of molinate with CYP116B1. Elution was monitored for absorbance at wavelengths of 254 and 220 nm and for signal on the mass spectrometric detector (MSD). A peak occurs in each of the traces at ~4 minutes.

The mass spectrum for the peak at ~4 minutes (Figure 5.17) shows a large peak at m/z 188.2 which corresponds to molinate. The spectrum also corresponds closely to the corresponding mass spectrum for the control reaction. This suggests that no reaction has occurred and only the substrate is present in the extracted reaction product.

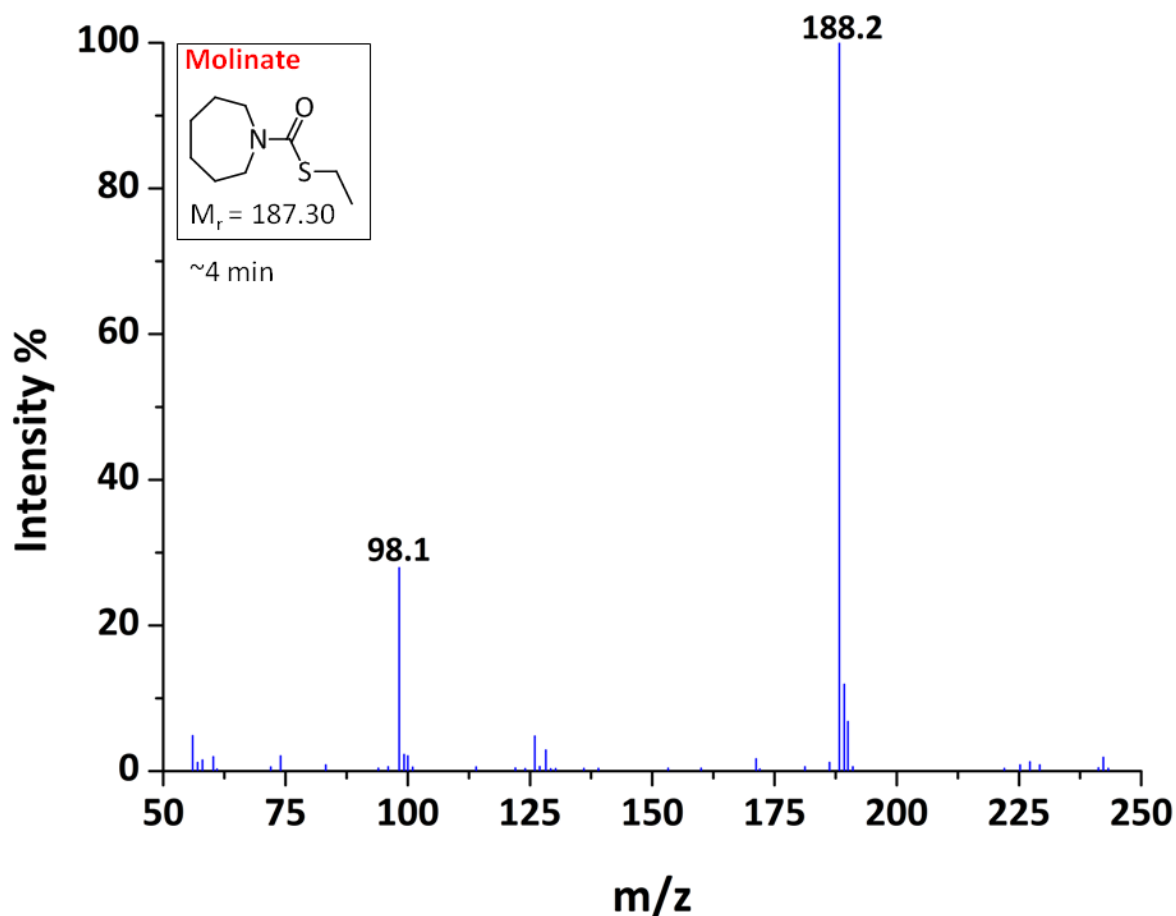


Figure 5.17. Mass spectrum of a CYP116B1 turnover reaction in the presence of substrate molinate. The mass spectrum was recorded for a fraction of the reaction with a HPLC retention time of ~4 minutes. The spectrum shows a peak at a m/z of 188.2 which corresponds (within 1 mass unit) to molinate. The equivalent mass spectrum for the control reaction is shown as an inset.

5.5.5. Prosulfocarb

The HPLC traces for the products of the reaction of prosulfocarb with CYP116B1 (Figure 5.18) show prominent elution peaks at ~1 and ~8 minutes. These elution peaks are visible in both absorbance traces and in the MSD trace.

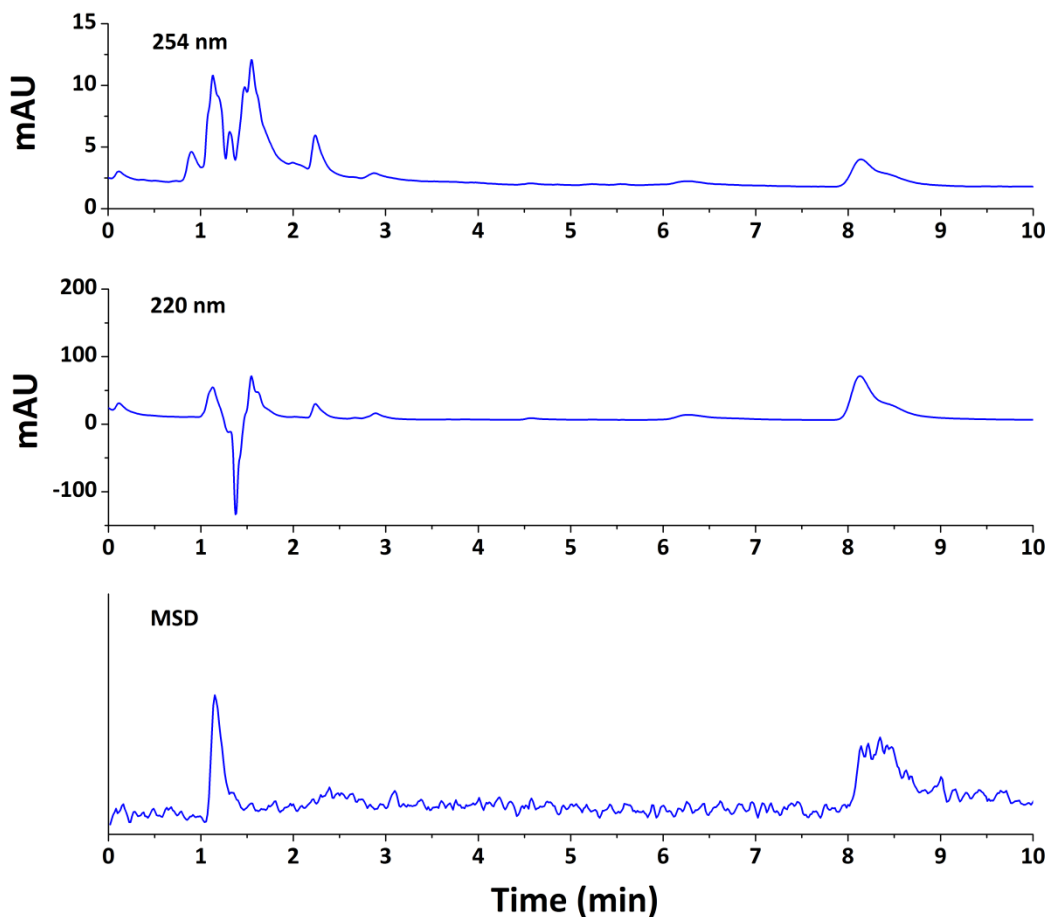


Figure 5.18. HPLC traces for the products of the reaction of prosulfocarb with CYP116B1. Elution was monitored for absorbance at wavelengths of 254 and 220 nm and for signal on the mass spectrometric detector (MSD). The largest peaks occur in both the 254 nm, 220 nm and MSD traces, and at ~1 and ~8 minutes.

The mass spectrum for the peak at ~8 minutes (Figure 5.19) shows a large peak at m/z 252.3 which corresponds to the molecular mass of prosulfocarb (251.39 g mol⁻¹). The spectrum also closely resembles the corresponding mass spectrum for the control reaction. This suggests that no reaction has occurred and only the substrate is present in the extracted reaction product.

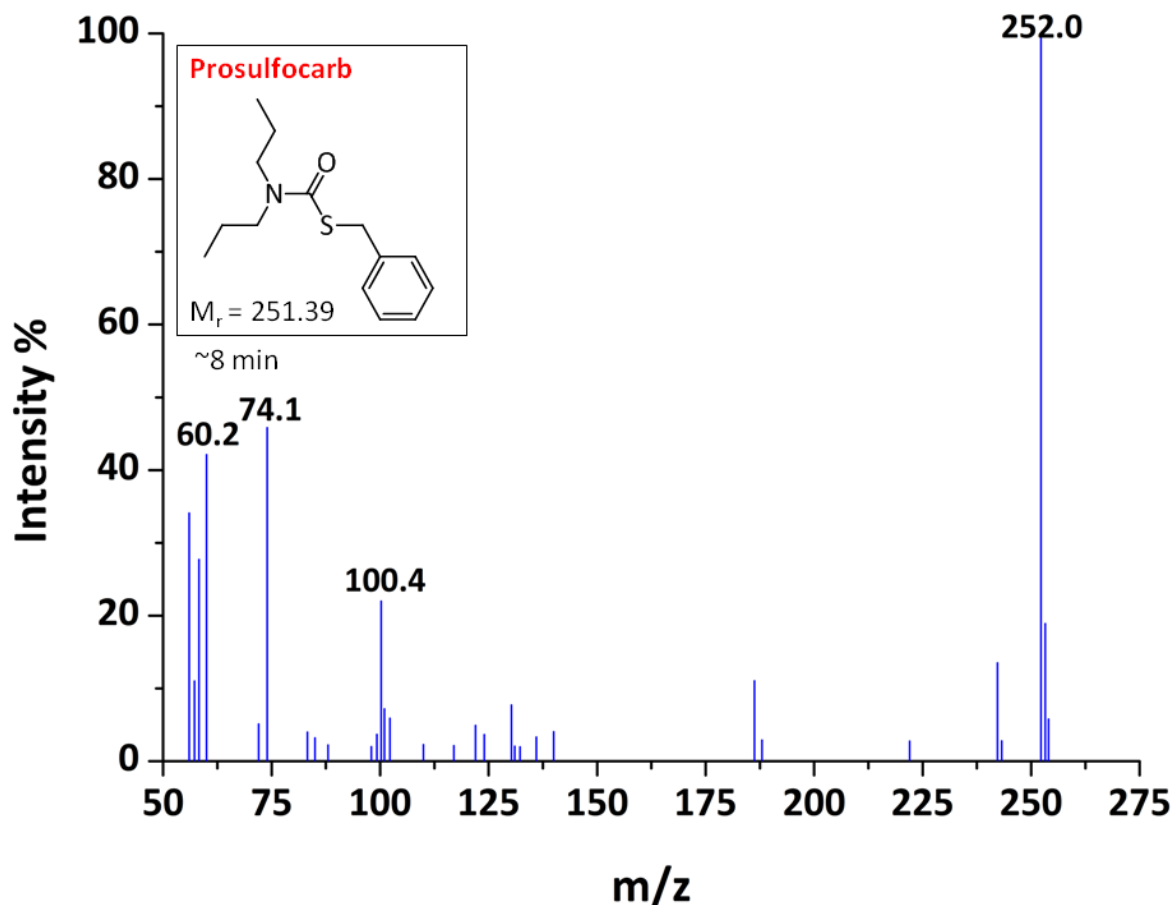


Figure 5.19. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate prosulfocarb. The mass spectrum was recorded for a fraction of the reaction with a HPLC retention time of ~8 minutes. The spectrum shows a peak at a m/z of 252.3 which corresponds (within 1 mass unit) to the molecular mass of prosulfocarb (251.39 g mol⁻¹). The equivalent mass spectrum for the control reaction is shown as an inset.

The mass spectrum for the peak at ~1 minute (Figure 5.20) also closely resembles the corresponding mass spectrum for the control reaction, although there is an increase in the relative intensity of the peak at m/z 122.1. The reason for this increase is unknown.

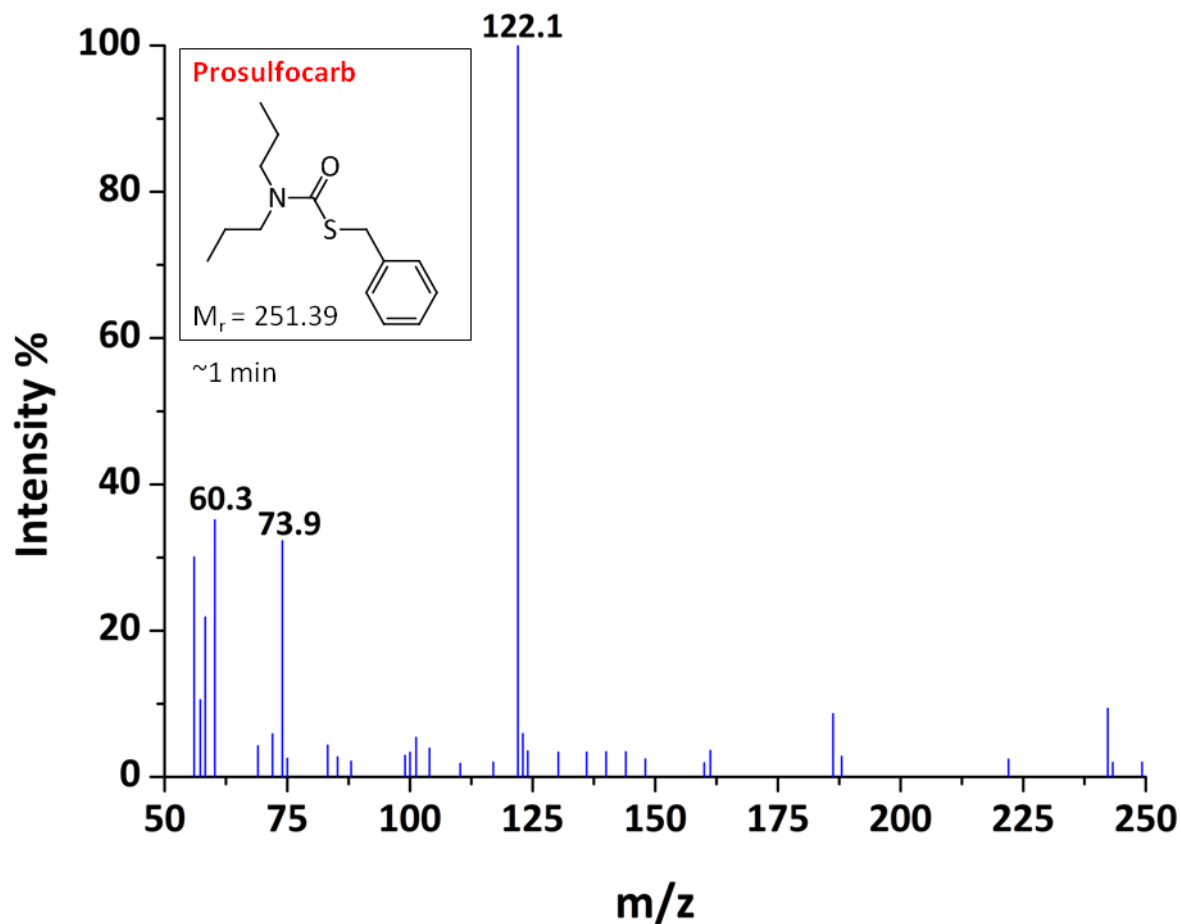


Figure 5.20. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate prosulfocarb. The mass spectrum was recorded for a fraction of the reaction with a HPLC retention time of ~1 minute. The spectrum shows an increased signal at m/z 122.1 compared to the corresponding mass spectrum for the control reaction.

5.5.6. Thiobencarb

The HPLC traces for the products of the reaction of thiobencarb with CYP116B1 (Figure 5.21) show prominent elution peaks at ~1 and ~6 minutes. The absorbance at ~1 minute is most intense in the 254 nm trace and that at ~5.4 minutes in the 220 nm trace. Only the elution peak at ~1 minute is also clearly visible in the MSD trace while that at ~5.4 minutes is less clear.

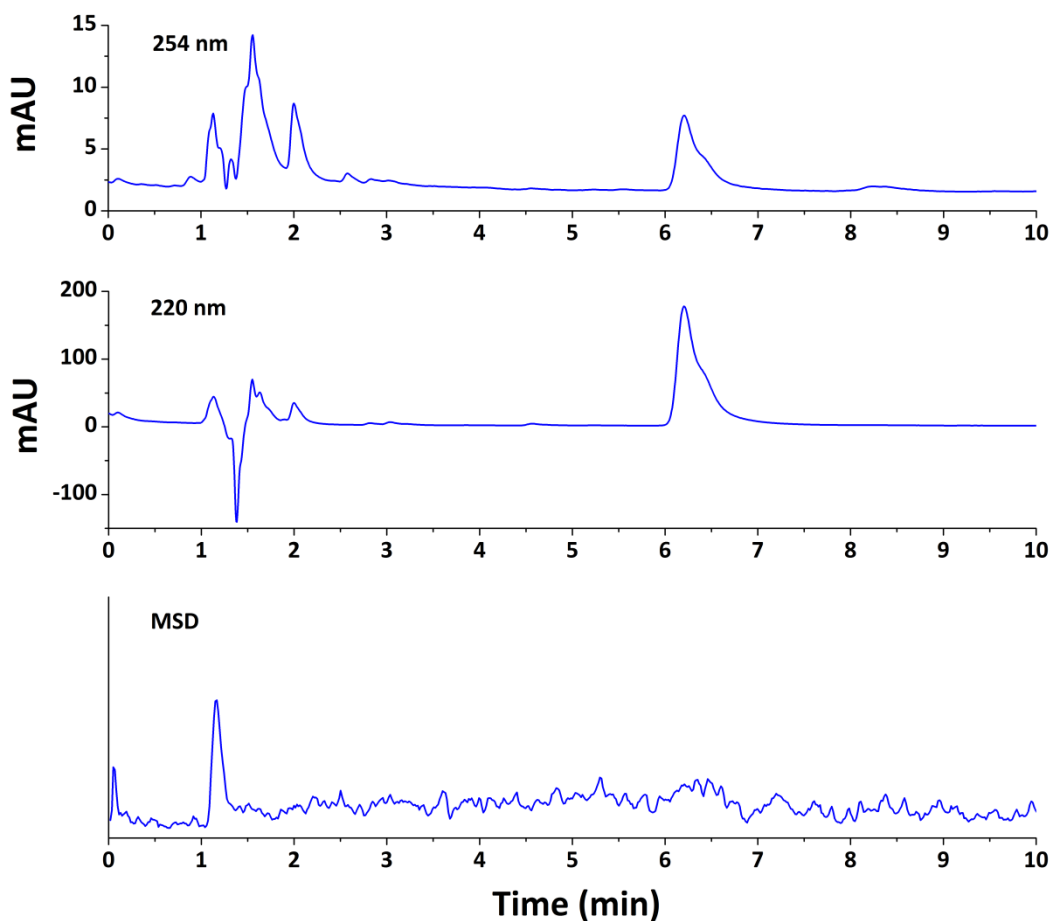


Figure 5.21. HPLC traces for the products of the reaction of thiobencarb with CYP116B1. Elution was monitored for absorbance at wavelengths of 254 and 220 nm and for signal on the mass spectrometric detector (MSD). The largest peaks occur in both the 254 nm and 220 nm traces at 1.66 minutes and in the MSD trace at 1.207 minutes.

The mass spectrum for the peak at ~1 minute (Figure 5.22) closely resembles the corresponding mass spectrum for the control reaction, although there is an increase in the relative intensity of the peak at m/z 122.1. The reason for this increase is unknown.

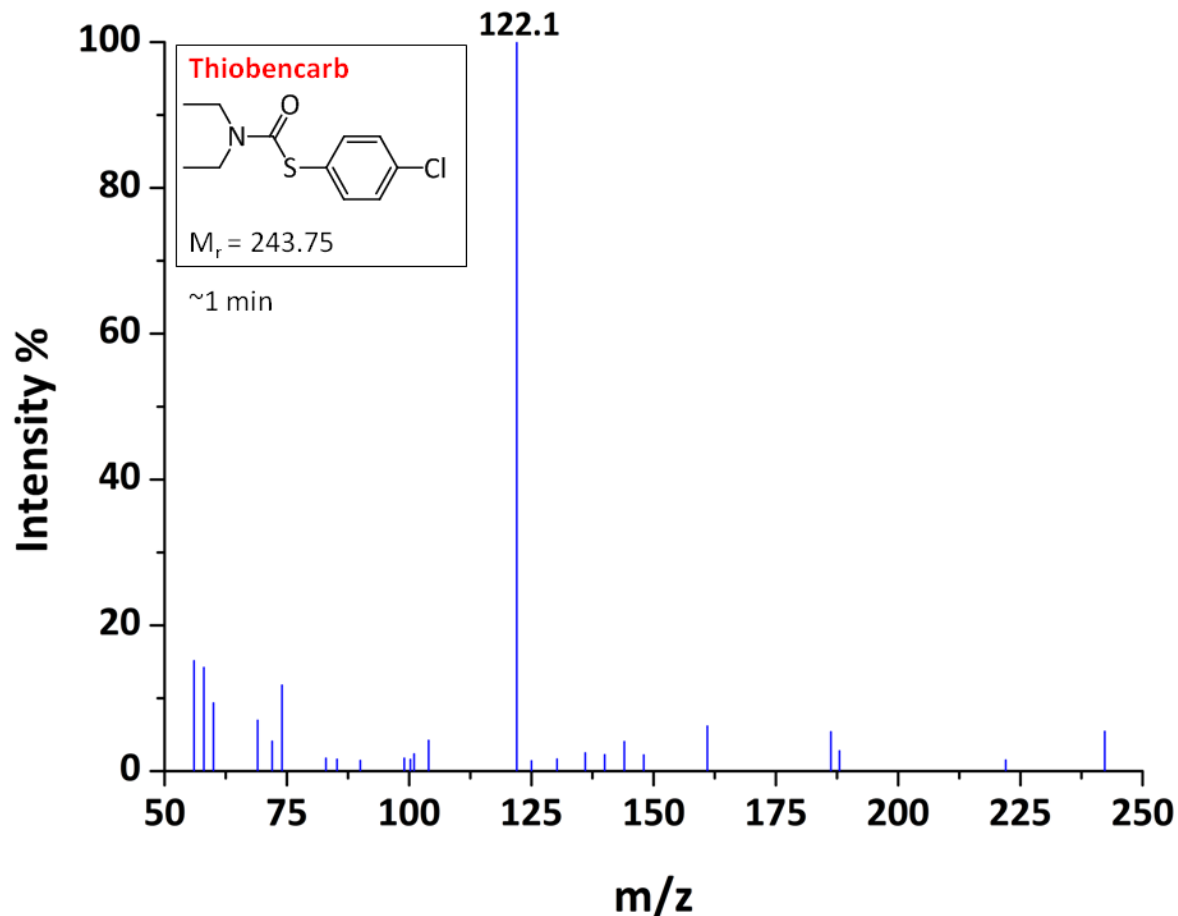


Figure 5.22. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate thiobencarb. The mass spectrum was recorded for a fraction of the reaction with a HPLC retention time of ~1 minute. The spectrum shows an increased signal at m/z 122.1 compared to the corresponding mass spectrum for the control reaction.

The mass spectrum for the peak at ~5.4 minutes (Figure 5.23) shows a large peak at m/z 242.3 which corresponds to the molecular mass of thiobencarb ($243.75 \text{ g mol}^{-1}$). The spectrum also closely resembles the corresponding mass spectrum for the control reaction. This suggests that no reaction has occurred and only the substrate is present in the extracted reaction product.

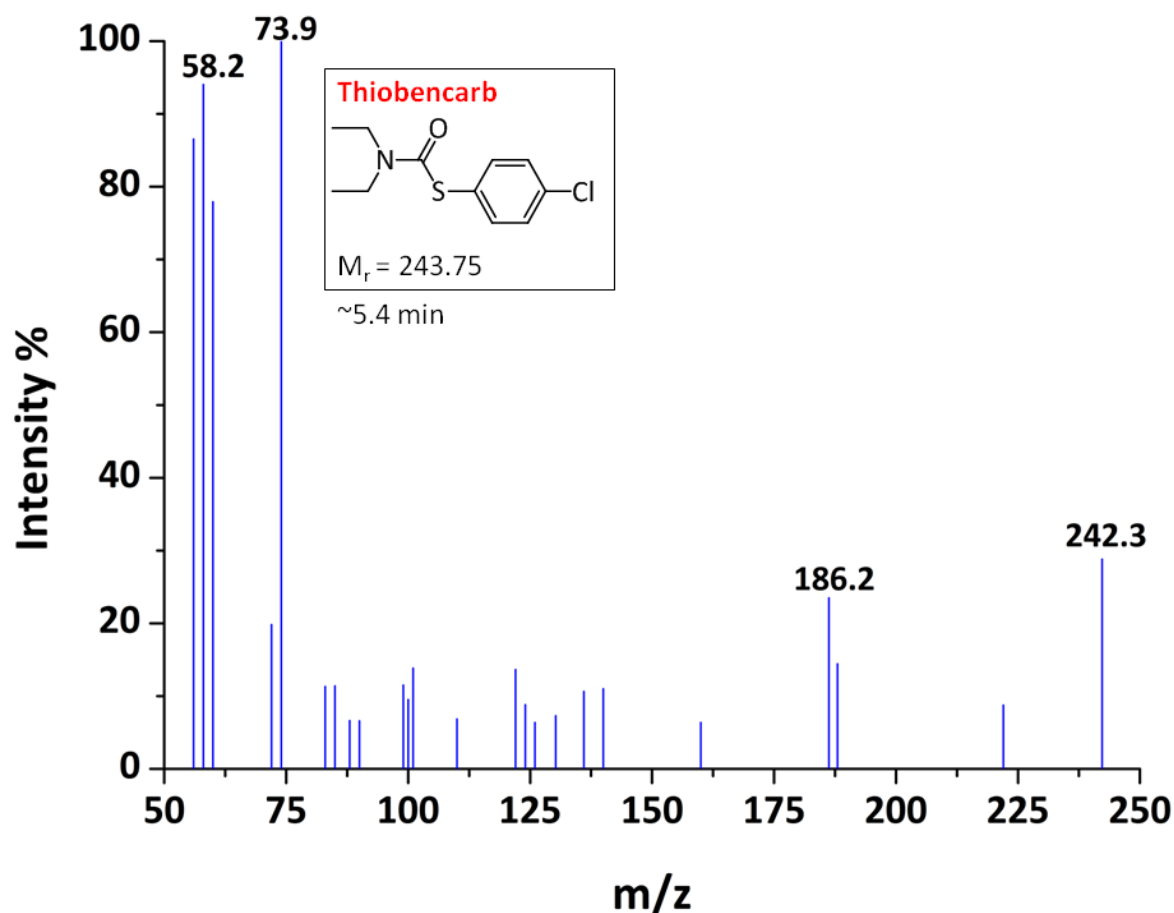


Figure 5.23. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate thiobencarb. The mass spectrum was recorded for a fraction of the reaction with a HPLC retention time of ~ 5.4 minutes. The spectrum shows a peak at m/z 242.3 which corresponds (within 1 mass unit) to the molecular mass of thiobencarb ($243.75 \text{ g mol}^{-1}$). The spectrum is also practically identical to the corresponding mass spectrum for the control reaction (presented as an inset).

5.5.7. Vanillic acid

The HPLC traces for the products of the reaction of vanillic acid with CYP116B1 (Figure 5.24) show prominent elution peaks at ~1 and ~1.5 minutes. The absorbance at ~1.5 minutes is clear in both the 254 nm and 220 nm traces. Only the elution peak at ~1 minute is discernable in the MSD trace.

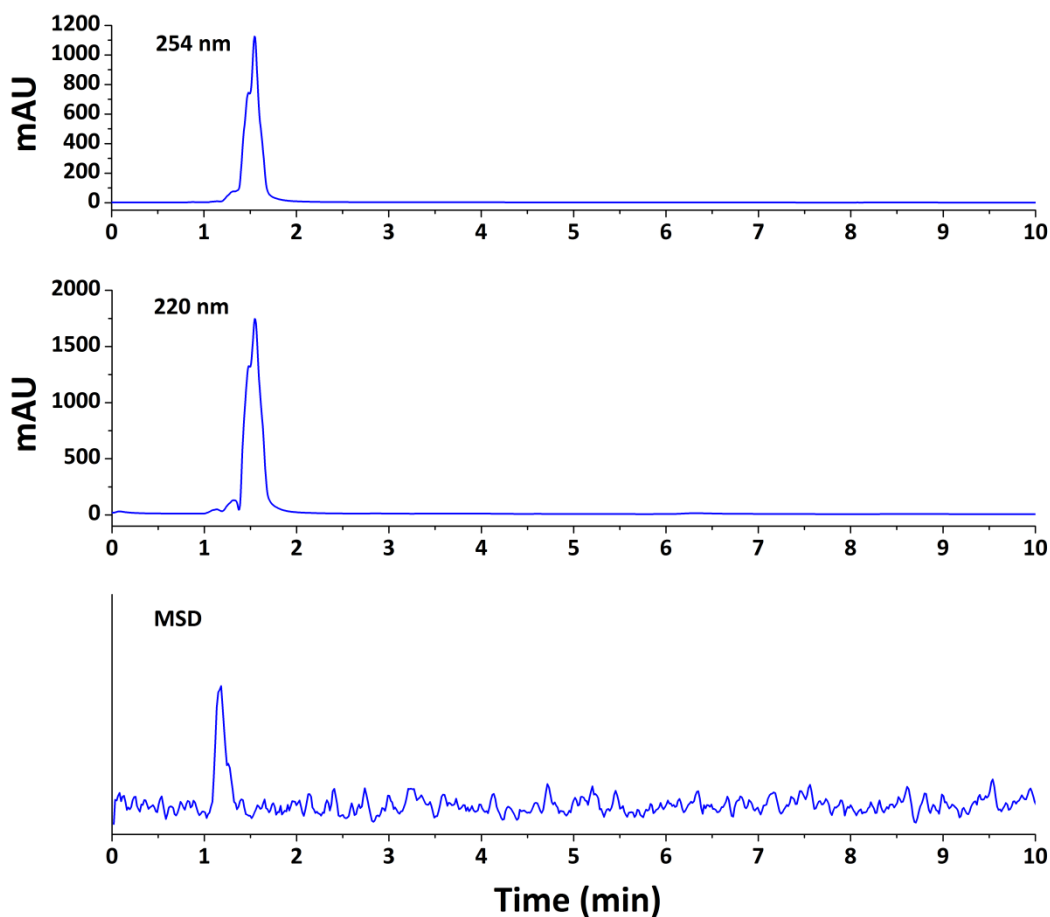


Figure 5.24. HPLC traces for the products of the reaction of vanillic acid with CYP116B1. Elution was monitored for absorbance at wavelengths of 254 and 220 nm and for signal on the mass spectrometric detector (MSD). The largest peak occurs in both the 254 nm and 220 nm traces at ~1.5 minutes, while the only peak in the MSD trace occurs at ~1 minute.

The mass spectrum for the peak at ~1 minute (Figure 5.25) closely resembles the corresponding mass spectrum for the control reaction. This suggests that no reaction has occurred and only the substrate is present in the extracted reaction product.

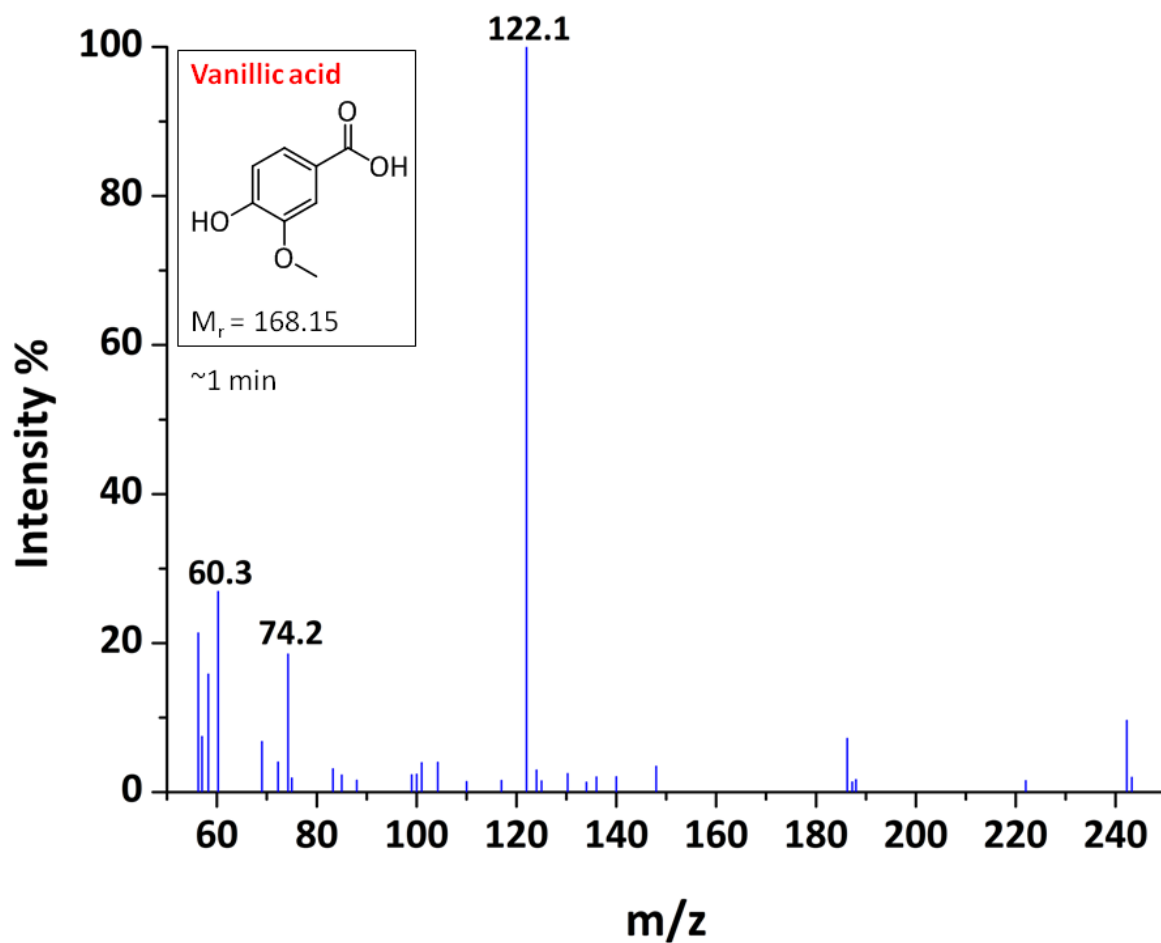


Figure 5.25. Mass spectrum of a CYP116B1 turnover reaction in the presence of the substrate candidate vanillic acid. The mass spectrum was recorded for a fraction of the reaction with a HPLC retention time of ~1 minute.

5.5.8. Vernolate

The HPLC traces for the products of the reaction of vernolate with CYP116B1 (Figure 5.26) show elution peaks at ~2 and ~7 minutes. The absorbance at ~2 minutes is most intense in the 254 nm trace, and that at ~7 minutes is most intense in the 220 nm trace. Both elution peaks are visible in the MSD trace.

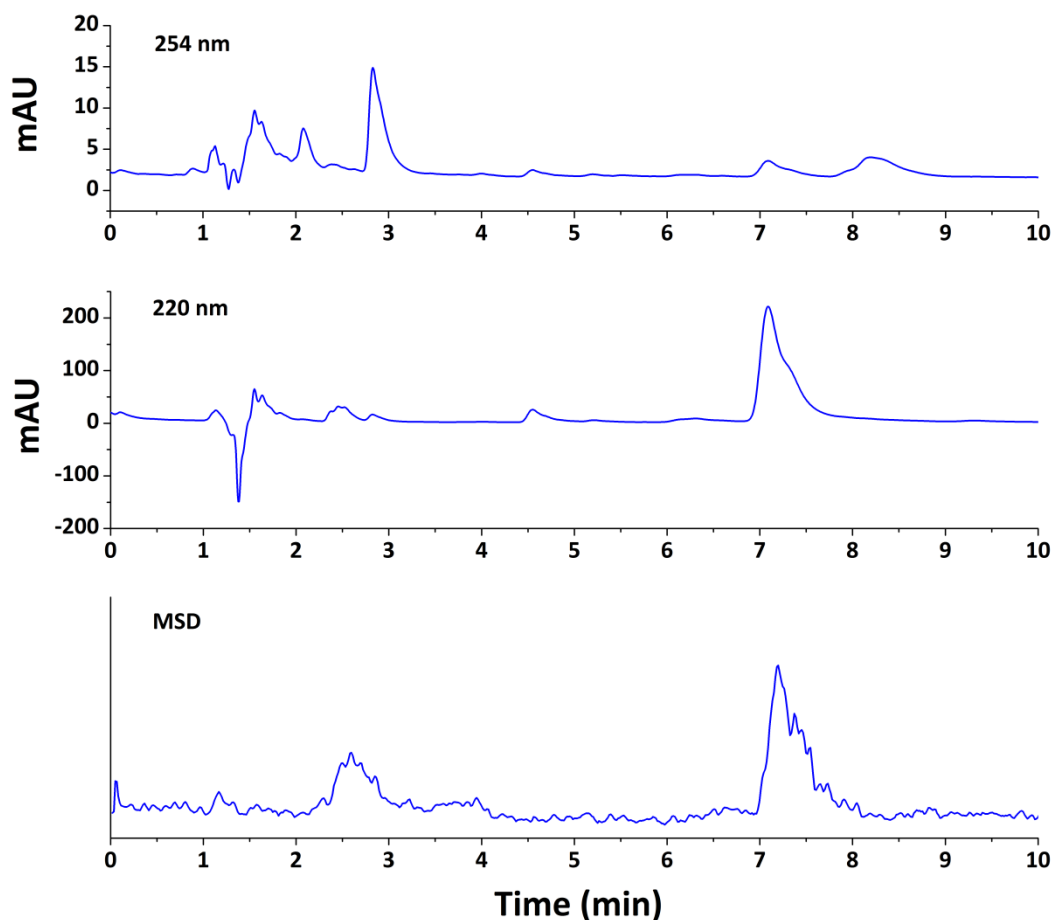


Figure 5.26. HPLC traces for the products of the reaction of atrazine with CYP116B1. Elution was monitored for absorbance at wavelengths of 254 and 220 nm and for signal on the mass spectrometric detector (MSD). The largest peaks occur in the 254 nm trace at 2.030 minutes and in the 220 nm trace at 7.069 minutes, with corresponding peaks in the MSD trace.

The mass spectrum for the peak at ~ 2.6 minutes shows a large peak at m/z 220.3 which corresponds to the molecular mass of hydroxy vernolate ($219.34 \text{ g mol}^{-1}$), a product that has previously been observed for the reaction of CYP116B1 with this substrate.

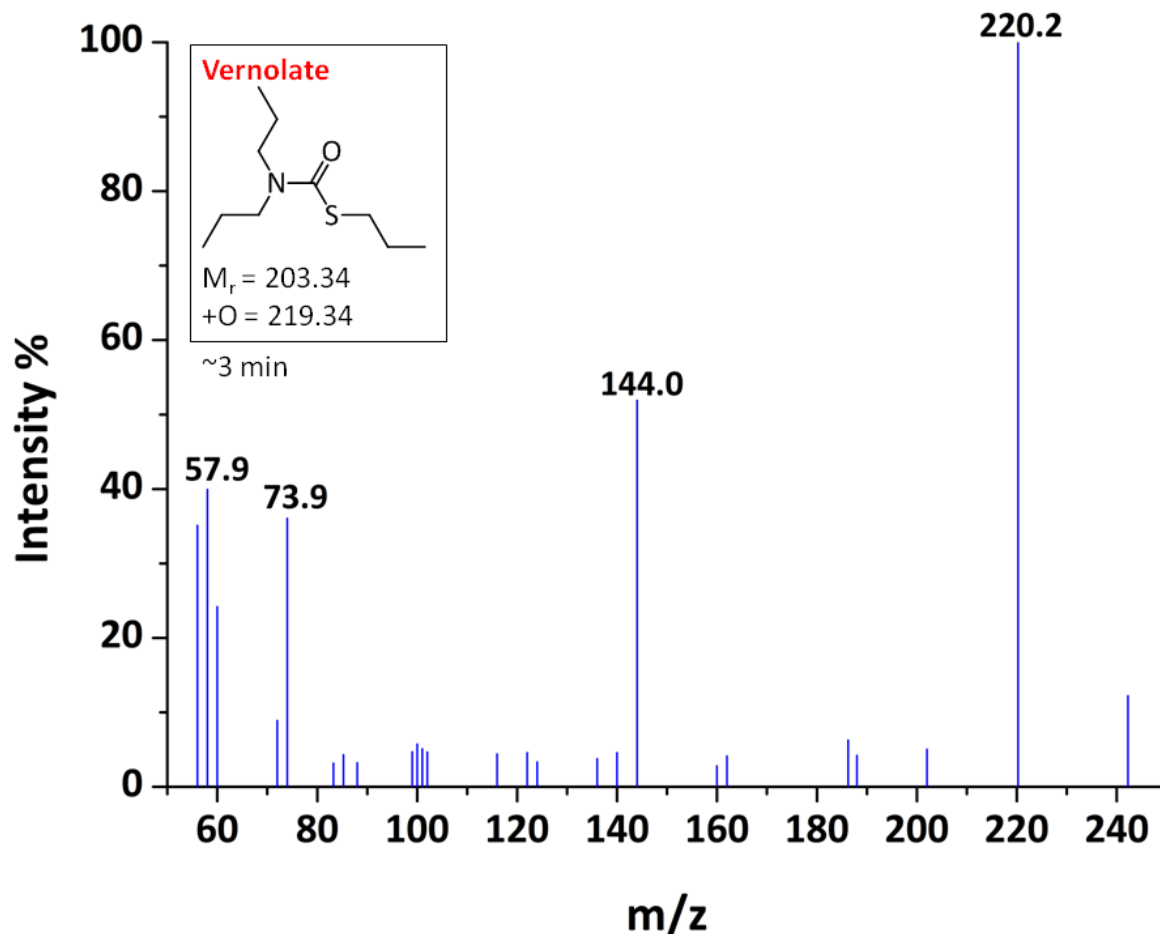


Figure 5.27. Mass spectrum of a CYP116B1 turnover reaction against substrate vernolate. The mass spectrum was recorded for a fraction of the reaction with a HPLC retention time of ~ 2.6 minutes. The spectrum shows a peak at a m/z ratio of 220.3 which corresponds (within 1 mass unit) to a hydroxylated product.

The peaks in the above mass spectrum at m/z ratios of 144.0 and 58.0 are the same as those seen in the mass spectrum for the reaction of CYP116B1 with EPTC. In that spectrum these peaks are assigned to *N*-depropyl EPTC ($M_r = 147.24 \text{ g mol}^{-1}$), and propanal ($M_r = 58.08 \text{ g mol}^{-1}$), two proposed products of the decomposition of hydroxy EPTC (see section 5.5.3). However, the molecular mass of *N*-depropyl vernolate is $161.27 \text{ g mol}^{-1}$. While this does correspond to a peak in the above mass spectrum at m/z 162.0, the relative intensity of this peak is low. This may cast doubt on the assignment of the peak at m/z 144.0 in the EPTC spectrum.

For EPTC and vernolate the product extract showed a species with a retention time of ~5 and ~3 minutes, respectively, whose mass spectra showed a signal at m/z corresponding within (one mass unit) to a hydroxylated product (with single ionisation). Surprisingly, for the other four substrate candidates that produced an enhancement of the rate of NADPH oxidation by CYP116B1, no reaction products could be identified from the HPLC-MS experiment. In these cases, peaks corresponding to the masses of the substrate candidates were identified, suggesting that turnover of the substrate candidates (*i.e.* hydroxylation or other reaction) had not occurred. It seems therefore, that these compounds stimulate the rate of oxidation of NADPH by CYP116B1 without themselves being turned-over. This may be effected by binding in the active site of the enzyme in a substrate-like manner allowing the flow of electrons through the haem but without becoming hydroxylated. The enzyme's catalytic cycle would then become one of futile NADPH oxidation followed by the release of peroxide or another oxygen species. This result is disappointing, as it means no true novel CYP116B1 substrates have been identified, but also intriguing as to the exact mechanism of the effect of these compounds upon the enzyme.

5.6. Steady State Kinetics of RDCYP116B1

Steady state kinetic experiments were conducted in order to examine the electron throughput of the isolated reductase domain and the theoretical limit imposed on the overall reaction of the CYP116B1 enzyme reaction by this domain. That is, by using artificial electron acceptors that rapidly remove electrons from the flavin or iron sulphur centres, it should be feasible to establish how fast electron transfer from NAD(P)H through these centres can occur. The physiological electron donor NAD(P)H was used as an electron source, along with a variety of non-physiological electron acceptors. The experiments measured the ability of RDCYP116B1 to transfer electrons from NAD(P)H through to these electron acceptors. The electron acceptors used were potassium ferricyanide (ferricyanide, $[\text{Fe}(\text{CN})_6]$), 2,6-dichlorophenolindophenol (DCPIP) and cytochrome *c*. This range of acceptors is chosen to be diagnostic of different aspects of the RDCYP116B1 reaction. $[\text{Fe}(\text{CN})_6]$ is comparatively small and so is able to easily penetrate the active site(s) of the enzyme and should gain access to both FMN and $[2\text{Fe-2S}]$ sites. $[\text{Fe}(\text{CN})_6]$ also carries a net

charge of -3. The reaction with this acceptor effectively defines a theoretical maximal rate for electron throughput in RDCYP116B1 (assuming that the electron transfer rate to this acceptor is faster than that to the haem domain of the enzyme, which is typically the case in P450 BM3 and in eukaryotic P450 reductase fusion enzymes). However, this will not be representative of the physiological reaction, and steady-state studies with *e.g.* EPTC and vernolate show that the reaction at the haem site with these substrates is quite slow (see Table 5.3).

DCPIP is a larger molecule which will only accept two electrons at a time, and therefore the reaction is almost certainly diagnostic of the capacity of the FMN of RDCYP116B1 (the only available two-electron donor, receiving its electrons as a hydride ion from NAD(P)H, to transfer electrons). Cytochrome *c* is a large, positively charged, haem containing, enzyme which could be taken to model the physiological electron acceptor of RDCYP116B1 *i.e.* its haem (P450) domain - HDCYP116B1. Cytochrome *c* is readily reduced by other P450 redox partners including CPR, and in such cases takes electrons almost exclusively from the FMN domain in this enzyme[210]. Thus, it is expected here that reduction of cytochrome *c* will be mediated mainly *via* electron transfer from the [2Fe-2S] domain of RDCYP116B1.

Initially, kinetic assays were carried out to determine the dependency on electron acceptor concentration of the RDCYP116B1 reaction. Assays were carried out by measuring changes in the visible absorbance spectrum specific to each acceptor upon its reduction. For $[\text{Fe}(\text{CN})_6]$ a decrease in absorbance at 420 nm ($\epsilon = 1.21 \text{ mM}^{-1}\text{cm}^{-1}$), for DCPIP a decrease in absorbance at 600 nm ($\epsilon = 21 \text{ mM}^{-1}\text{cm}^{-1}$), and for cytochrome *c* an increase in absorbance at 550 nm ($\epsilon = 22.6 \text{ mM}^{-1}\text{cm}^{-1}$). Kinetic assays were performed at a range of electron acceptor concentrations, using a UV/Vis. spectrophotometer and taking the initial rate (full assay protocol is described in section 2.4.16). Assays were performed in triplicate and the mean values of the initial rate were plotted against the electron acceptor concentration. The resultant kinetic curves were fitted using the Michaelis-Menten equation (section 1.2.1, Equation 1.1) and are presented in Figure 5.28.

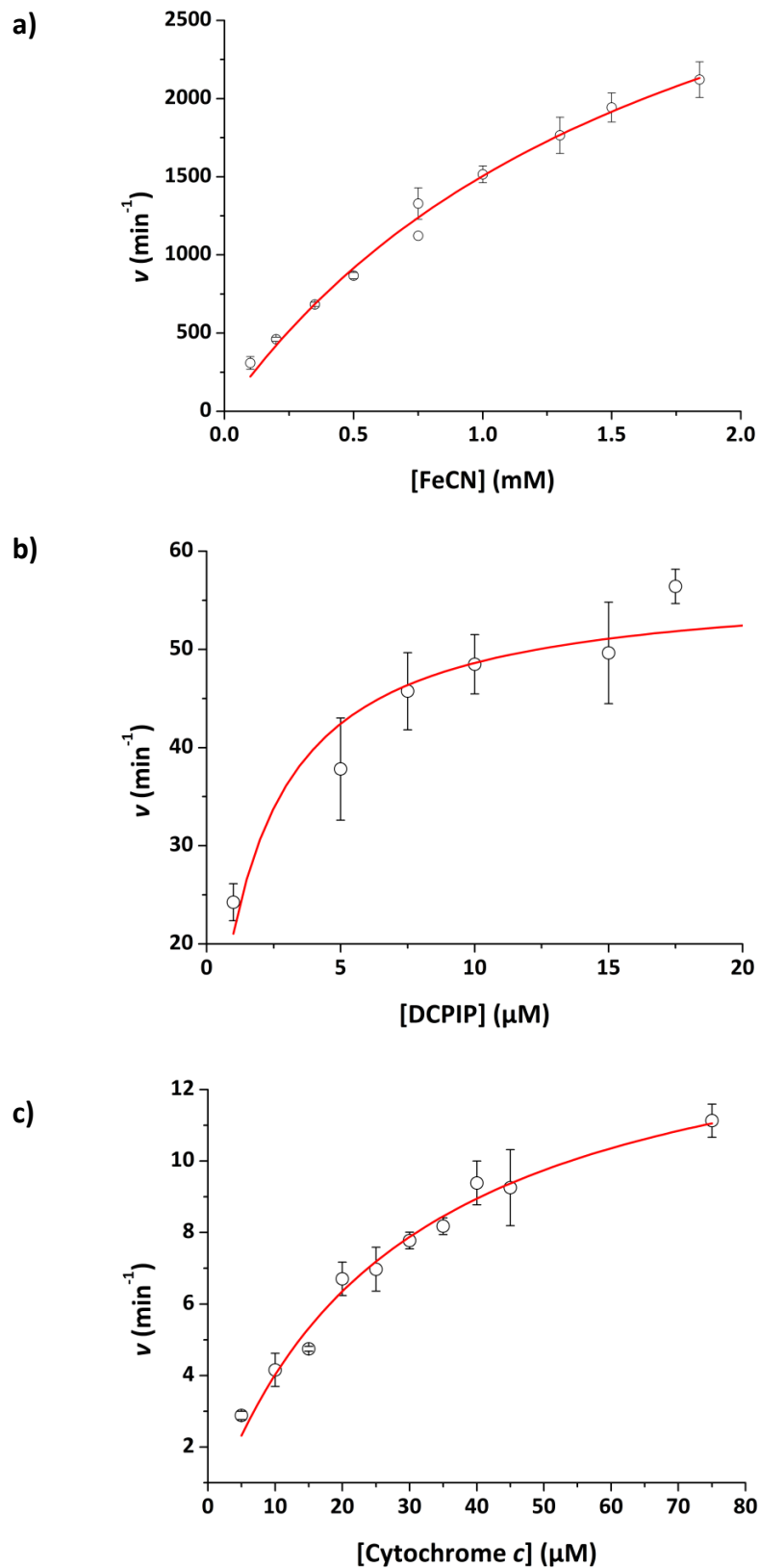


Figure 5.28. Kinetic curves showing the dependency of the rate of NADPH oxidation on the concentration of electron acceptors, in assays using RDCYP116B1. Assays were carried out with a range of concentrations of electron acceptors. Those used were a) ferricyanide $[\text{Fe}(\text{CN})_6]$ (0.1 – 1.8 mM), b) 2,6-dichlorophenolindophenol (DCPIP) (1 – 17.5 μM) and c) cytochrome *c* (5 – 75 μM). All assays were carried out in the presence of NADPH (200 μM). Values of the associated kinetic constants K_M and k_{cat} are presented in Table 5.4.

The fits for $[\text{Fe}(\text{CN})_6]$ and cytochrome *c* project the maximal rate at a rather higher concentration of the electron acceptor than was used in the experiment. In the case of ferricyanide, this reflects the relatively weak binding (low affinity) of this polar molecule for the enzyme, along with the fact that (at 2 mM) absorbance contributions from the substrate became large and it was not feasible to obtain accurate measurements at higher substrate concentrations. Nonetheless, a reasonably accurate estimate of k_{cat} and K_{m} was obtained (see Table 5.4). In the case of cytochrome *c*, it is clear that saturation with the electron acceptor is near-complete, and this is also reflected in the much lower error on the K_{m} and k_{cat} parameters (compared with ferricyanide) for this substrate (see Table 5.4).

Electron acceptor	k_{cat} (min^{-1})	K_{M} (μM)
Ferricyanide	4200 \pm 400	1800 \pm 300
DCPIP	57 \pm 4	2 \pm 0.5
Cytochrome <i>c</i>	15 \pm 1	30 \pm 4

Table 5.4. Kinetic constants (k_{cat} and K_{M}) for the RDCYP116B1 reaction showing the dependency on the concentration of three electron acceptors. The values are derived from fitting of kinetic curves to the Michaelis-Menten equation.

The results tabulated above show that, as expected, the fastest rate constant for NADPH consumption is achieved with $[\text{Fe}(\text{CN})_6]$, with lower rate constants for DCPIP and cytochrome *c* reduction. It should also be remembered that (since $[\text{Fe}(\text{CN})_6]$ and cytochrome *c* are one electron acceptors, whereas DCPIP is a two-electron acceptor) that the rates of product reduction are actually two-fold that of NADPH oxidation in the cases of FeCN and cytochrome *c*. The data for ferricyanide are probably indicative of the fact that $[\text{Fe}(\text{CN})_6]$ is a small molecule that can gain better access to the redox centres in RDCYP116B1 than can the other electron acceptors. For this reason it is also suggested, by the high rates of NADPH oxidation for this electron acceptor, that it receives electrons directly from the FMN cofactor of RDCYP116B1, thus bypassing the $[\text{2Fe-2S}]$ cluster. This rate would therefore be reporting principally on the transfer of electrons between NADPH and the FMN cofactor of RDCYP116B1 and thus describe a lower limit for the rate of hydride transfer between NAD(P)H and the FMN at 2700 min^{-1} this suggests that the rate of hydride transfer under the conditions in this experiment is $> 45 \text{ s}^{-1}$.

DCPIP, which requires the concomitant transfer of two electrons, shows a much slower rate than $[\text{Fe}(\text{CN})_6]$. This is probably a consequence of lower efficiency of two-electron transfer from the FMN and might also reflect the occlusion of the binding site to this larger molecule.

Cytochrome *c* is arguably the most similar of the electron acceptors to the natural acceptor (HDCYP116B1), although its protein fold is rather different and it is a rather smaller protein. In the case of CPR, electrons are transferred from NAD(P)H by the FAD cofactor to electron acceptors, including cytochrome *c*, via the FMN cofactor[211]. By analogy, cytochrome *c* may only be able to receive electrons through the [2Fe-2S] cluster of the ferredoxin-like subdomain of RDCYP116B1. If this is the case, then the limiting rate for this acceptor ($k_{\text{cat}} = 15 \pm 1 \text{ min}^{-1}$) puts a lower limit on the rate of electron transfer between the FMN and [2Fe-2S] centres in RDCYP116B1 and HDCYP116B1, although it is probably the case that the inter-cofactor electron transfer occurs rather faster than this. Indeed, for PDOR, the rate of electron transfer between FMN and [2Fe-2S] cofactors has been experimentally measured, and is significantly faster ($k = 35 \text{ s}^{-1}$)[131].

Further assays were carried out to probe the kinetic dependency of electron transfer by RDCYP116B1 upon NAD(P)H concentration. These experiments were carried out for each of the three electron acceptors, at fixed concentrations (1 mM $[\text{Fe}(\text{CN})_6]$, 20 μM DCPIP or 45 μM cytochrome *c*). The activity was measured by observing the same changes in the visible absorbance spectrum, for each electron acceptor, as detailed above. The experiment was carried out at a range of NADH/NADPH concentrations (typically 0.1 - 2 mM). These experiments show that at saturating NAD(P)H concentrations the maximal rate of NADPH oxidation (represented by the data for ferricyanide) is $>240 \text{ s}^{-1}$, and that the lower limit for FMN to [2Fe-2S] electron transfer (represented by the cytochrome *c* data) is 43 min^{-1} . The kinetic constants associated with these experiments are presented in Table 5.5.

Electron Acceptor	NADPH		NADH	
	k_{cat} (min^{-1})	K_{M} (μM)	k_{cat} (min^{-1})	K_{M} (μM)
Ferricyanide	14600 \pm 3300	1300 \pm 450	15900 \pm 2900	1400 \pm 400
DCPIP	120 \pm 23	480 \pm 190	93 \pm 9	860 \pm 130
Cytochrome c	43 \pm 10	1.6 \pm 0.6	~40	~2

Table 5.5. Kinetic constants (k_{cat} and K_{M}) for the RDCYP116B1 reaction showing the dependency on the concentration of electron donor (NADH or NADPH). Assays were carried out in the presence of fixed concentrations of electron acceptor molecules: Potassium ferricyanide (1000 μM), DCPIP (20 μM) and cytochrome *c* (45 μM). A range of concentrations of two-electron donors NADH and NADPH were used, typically 0.1 – 2 mM.

The data in the table above show that RDCYP116B1 consistently shows a slight preference for NADPH. This is evidenced by the K_{M} values for NADPH being lower than those for NADH for DCPIP and cytochrome *c*. Additionally, for these two electron acceptors, a slightly higher value of k_{cat} is observed with NADPH compared to NADH. In the case of ferricyanide there is no real difference in the k_{cat} or K_{M} values within the margins of error. The results for cytochrome *c* and DCPIP agree with work on CYP116B2 (P450RhF) which shows a preference for NADPH over NADH binding[190]. In both cases this is the opposite coenzyme preference to that shown by PDOR, which has been found to favour NADH binding[132].

The suggestion is that a change has occurred during the evolution of this domain, from PDOR or common ancestors, which has switched the coenzyme specificity to NADPH. In the ferredoxin reductase enzyme family (FNR), of which PDOR is a member, a sequence of two adjacent residues, Ser-234 & Arg-235 (see Figure 5.29), has been identified as the primary determinant of NADPH specificity[132]. These residues have been shown, by structural characterisation, to interact with the 2'-phosphate group of NADPH[212]. Upon comparison of their amino acid sequences, these residues, not present in PDOR, are aligned at the appropriate position in CYP116B1 as Ser-689 & Arg-690 (approximately Ser-233 & Arg-234 relative to the reductase domain start). In PDOR, it is thought that another residue (Asp-173) helps to bind NADH by interacting with its ribosyl chain[132] (see Figure 5.30).

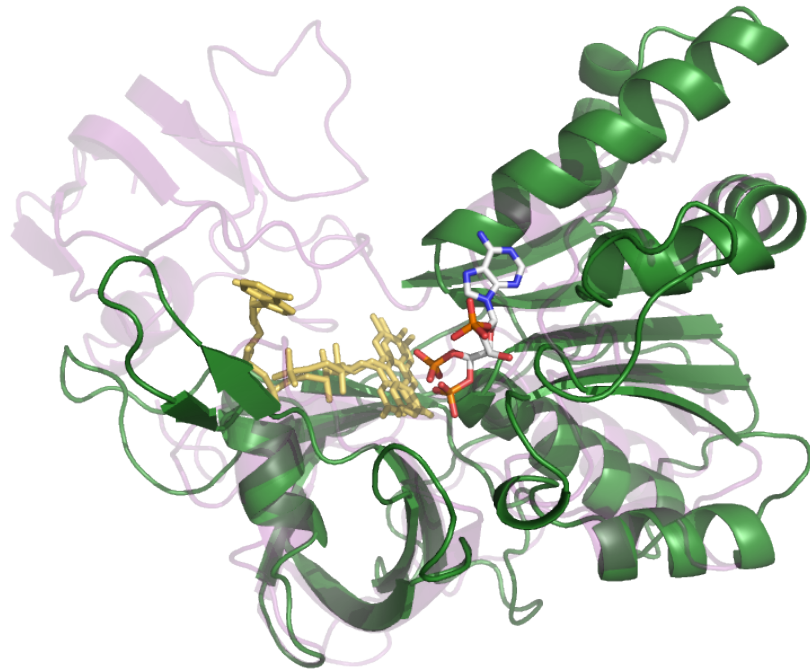


Figure 5.29. Images of the aligned structures of spinach ferredoxin reductase with bound 2'-phospho-5'-adenosine monophosphate (AMP) (RCSB PDB accession code: 1FND) (green); and PDOR from *Pseudomonas cepacia* (accession code 2PIA) (purple). Both structures are shown as cartoons with their flavin cofactors in yellow. The bound 2'-phospho-5'-AMP from the FNR structure is also shown.

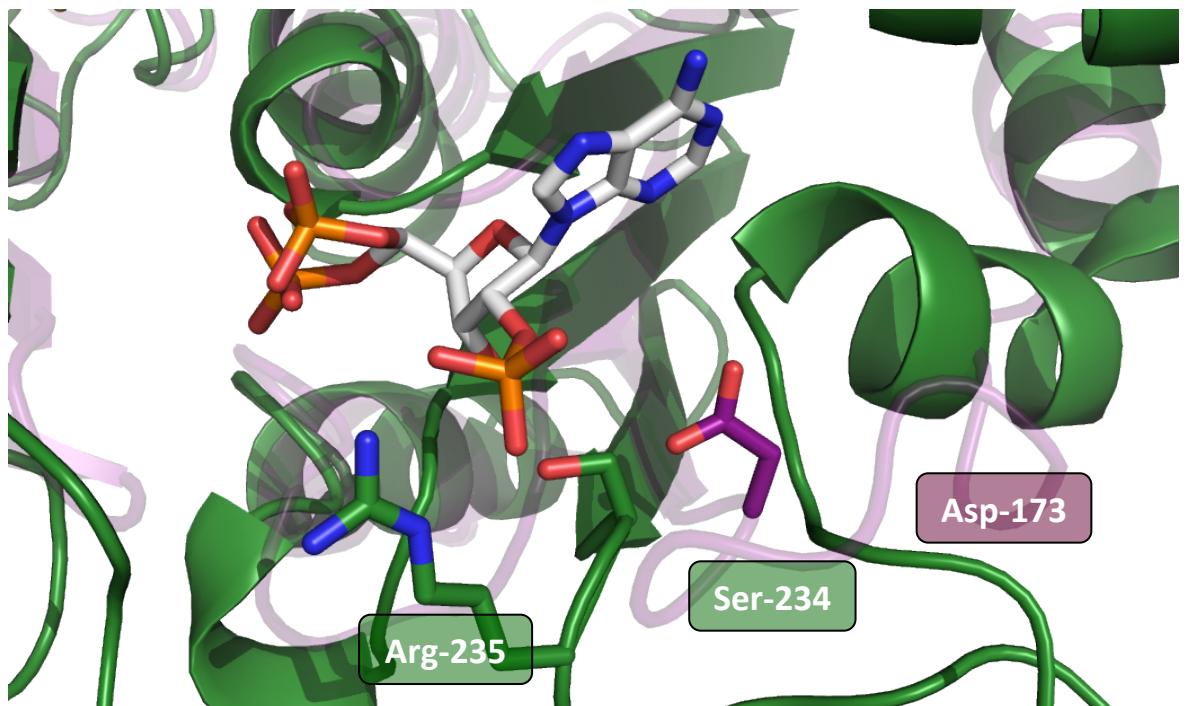


Figure 5.30. Close-up view of the structures shown in Figure 5.29. From the FNR structure, residues Ser-234 & Arg-235, thought to interact with the 2'- phosphate of NADPH, are shown. From the PDOR structure Asp-173, thought to interact with the ribityl chain of NADH, is shown.

5.7. Summary & Discussion

In this chapter, the work has focussed on how CYP116B1, a fusion protein consisting of P450 haem and PDOR-like reductase domains, functions as an enzyme. The work began with the construction of a substrate library composed of identified substrates of CYP116B1, compounds related to these and the substrates of evolutionarily related enzymes. From this library a panel of eight substrate candidates was selected to provide a representative sample of the entire library.

Binding titrations, carried out on HDCYP116B1 using existing substrates, revealed that the effect of binding interactions on haem ligation and spin state is limited. This result was supported by redox potentiometry which showed that a small magnitude of shift in the HDCYP116B1 haem-iron redox potential is effected by substrate binding. This experiment also suggested, at least for the isolated haem domain, that CYP116B1 haem-iron redox potentials are substantially more positive than the majority other P450 enzymes studied.

The binding of inhibitors to HDCYP116B1 yielded some unusual results which explained previous observations regarding the retention of imidazole from the purification process by the haem of HDCYP116B1 and FLCYP116B1 proteins (see Chapter 4). Imidazole was found to bind to HDCYP116B1 in two phases with different dependencies on concentration of imidazole. The first was a tightly bound population showing a hyperbolic binding dependency, and a second more weakly bound population showing a sigmoidal dependency. The initial, tightly bound, phase could help to explain the residual imidazole from the purification process of HDCYP116B1 and FLCYP116B1 proteins. The imidazole derivative inhibitor 2-phenylimidazole also showed relatively tight binding to HDCYP116B1, compared to many other P450 enzymes, with a K_d value of $72 \pm 5 \mu\text{M}$. This result may suggest a more open substrate binding cavity around the haem allowing a more favourable binding mode for this ligand than in the other enzymes.

Redox potentiometry showed that the HDCYP116B1 haem iron redox potential is significantly more positive than many other P450 enzymes. This experiment also showed that the binding of substrate EPTC produces a very small shift in haem iron redox potential compared to cases of other P450 enzymes with their substrates.

Kinetic studies were carried out to determine whether stimulation of NADPH oxidation, by FLCYP116B1, could be effected by any of the substrate candidates, and to what extent. These experiments showed that for six of the eight substrate candidates, stimulation of NADPH oxidation was achieved. Kinetic constants (k_{cat} and $K_{\text{M}}/K_{\text{H}}$) were also determined for these reactions (see Table 5.3) revealing the rate constant for NADPH oxidation induced by each substrate candidate. Most of the substrate candidates achieved similar rate constants of NADPH oxidation with k_{cat} values between 1-5 min^{-1} . The extremes were occupied by prosulfocarb and vanillic acid whose k_{cat} values were around twenty times higher and one hundred times lower than any other substrate candidate, respectively. Three of the substrate candidates (di-allate, prosulfocarb and thiobencarb) produced sigmoidal kinetic dependencies, suggesting the same apparent binding cooperativity seen for HDCYP102A3 with various substrates (see Chapter 3).

Product determination experiments were carried out to investigate whether any products of the CYP116B1 reaction with the substrate candidates could be observed by HPLC-MS. In these experiments, only two of the substrate candidates showed peaks in their mass spectra confirming hydroxylated products (vernolate and EPTC).

An overview of the substrate candidates stimulating NADPH oxidation by CYP116B1, and those with identified products is shown in Table 5.6.

Substrate candidate	NADPH oxidation	Products
Atrazine	-	-
Di-allate	+	-
EPTC	+	+
Molinate	-	-
Prosulfocarb	+	-
Thiobencarb	+	-
Vanillic acid	+	-
Vernolate	+	+

Table 5.6. Collated data for substrate candidates showing whether NADPH oxidation was stimulated, and whether reaction products were identified, for each substrate candidate with FLCYP116B1.

The suggestion from Table 5.6 is that many of the substrate candidates stimulate the oxidation of NAD(P)H by the enzyme, but are not themselves oxidised. This may result in the futile cycling of electrons through the enzyme with no substrate oxidation and reduction of molecular oxygen occurring instead.

The kinetics of the reductase domain of CYP116B1 were also investigated with steady state experiments. A range of non-physiological electron acceptors, with varying properties, was used to probe various aspects of electron throughput in RDCYP116B1. A comparison can be drawn between the rates of substrate turnover for the full length enzyme (see section 5.4) and the lower limit defined for the rate of electron throughput in RDCYP116B1 for the same concentration of NADPH (represented by the rate of electron transfer between NADPH and cytochrome *c*, $15 \pm 1 \text{ min}^{-1}$). This comparison shows that electron transfer from NADPH is at least twice as fast as substrate turnover in most cases. This would suggest that the rate of the substrate oxidation reaction is not limited by the redox partner reaction and particularly not by the rate of FMN reduction by NADPH – which occurs at $>45 \text{ s}^{-1}$ based on measurements of ferricyanide reduction rates.

Following the work presented in this chapter, a number of further experiments are suggested. Further investigation of the unusual phenomena seen for the binding of imidazole to CYP116B1 would be warranted. Techniques such as resonance Raman spectroscopy or isothermal titration calorimetry could be employed to gain more

insight on this issue. Renewed attempts to determine redox potentials for CYP116B1 and its isolated domains could reveal more details of the unusually positive values recorded for the isolated haem domain. This work would first involve determining experimental conditions under which the FLCYP116B1 protein is stable and could then allow for analysis of the influence of the redox partner on the haem iron potential.

Continued work using the substrate candidates would be advised to confirm whether turnover does in fact occur in the presence of substrates other than EPTC and vernolate. This may involve further product determination experiments where the duration of reaction is a controlled variable, which would allow for the appearance of products to be monitored over time.

The confirmed activity of CYP116B1 against two thiocarbamate pesticides is encouraging. With further work to develop a better understanding of the enzyme, it may be possible to engineer it to develop activity against a range of other thiocarbamates and related molecules. This makes it a potentially valuable enzyme for the fields of bioremediation and synthetic chemistry.

Characterisation of Novel Cytochrome P450 Fusion Systems

Chapter 6

Summary and Bibliography

6. Summary and Bibliography

6.1. Summary

The work presented in this thesis concerns two cytochrome P450 fusion systems. These are enzymes where a P450 haem domain and another protein domain are expressed together as a single protein chain. In the case of the fusion enzymes studied, the fused domain is a redox partner enzyme involved in supplying electrons from NAD(P)H coenzymes to the P450 enzyme. The first of these is CYP102A3 from *Bacillus subtilis* which is an example of a P450-CPR fusion and has known activity as a fatty acid hydroxylase. This enzyme shares substantial homology and amino acid sequence identity with the well studied model enzyme CYP102A1, also known as BM3. The second system was CYP116B1, a P450 fusion with a PDOR-like reductase domain. The physiological function of this enzyme is not known, but previous work has demonstrated that hydroxylation of thiocarbamate herbicides EPTC and vernolate is carried out by this enzyme, and thus the enzyme is putatively involved in a detoxification reaction.

6.1.1. HDCYP102A3

A key aim was the structural characterisation of the haem domain of CYP102A3 by x-ray crystallography. Though this technique was unsuccessful, a collection of other experiments provided insights into structural aspects of the domain. EPR and resonance Raman spectroscopies showed that the environment around the haem cofactor of CYP102A3 is practically identical to that of BM3. Homology modelling provided an idea of what the structure of HDCYP102A3 might be, and this also showed very high similarity to the BM3 structure used as a template. Indeed, the homology model constructed shows very little visible variation when overlaid with the template structure of the haem domain of BM3. Contrasting with this close structural agreement is the observation, from previous work, that the substrate preference and regioselectivity of hydroxylation vary quite markedly between the two enzymes, although both are fatty acid hydroxylases. This is further demonstrated by the substrate binding work presented in this thesis that shows differences in the pattern of dissociation constants for HDCYP102A3 compared to both BM3 and the full-length CYP102A3 fusion. This includes the apparent cooperative binding behaviour exhibited by certain branched chain fatty acid substrates, which is not seen

for BM3. DSC and chemical denaturation together showed that the HDCYP102A3 protein is reasonably stable to denaturing conditions, at least with respect to environmental conditions that might be encountered by *Bacillus subtilis*. These techniques also supported a previous theory suggesting that, during the denaturing of BM3, the reductase domain unfolds first followed by the haem domain.

6.1.2. CYP116B1

The work on CYP116B1 included the development of an efficient regime for expression and purification of the full length enzyme and its isolated haem domain. These studies also revealed unusual imidazole binding properties of the CYP116B1 haem domain, which were further examined by EPR studies and optical binding titrations. This work showed an apparently biphasic binding profile for imidazole with HDCYP116B1, with two populations of bound imidazole, the first tightly bound ($K_d \sim 18 \mu\text{M}$) and the second less tightly bound ($K_H \sim 580 \mu\text{M}$). EPR studies of the full length CYP116B1 protein also confirmed the presence and stoichiometric binding of the [2Fe-2s] cofactor. Light scattering and DSC demonstrated that the CYP116B1 protein is stable across a range of temperatures and is largely monomeric in solution. This suggests that electron transfer in CYP116B1 occurs within the monomer, in contrast to the BM3 fusion enzyme (and thus likely also CYP102A3) which is catalytically active as a dimer with electron transfer occurring between monomers[73]. Optical titrations and redox potentiometry experiments showed that the binding of known substrates is quite inefficient, at least by normal measures of substrate binding in P450 enzymes, and in terms of affinity and lack of ability to shift the haem-iron spin-state equilibrium extensively towards high-spin. Kinetic and product determination studies revealed that while many molecules related to the known substrates stimulate the flow of electrons through CYP116B1, few of them actually undergo oxidation by the enzyme. Studies of the isolated reductase domain revealed that, at present, purification is unreliable for this isolated domain. Kinetic studies confirmed the preference of NADPH over NADH as an electron source. This demonstrates that while RDCYP116B1 is structurally similar to PDOR in its reductase domain, its coenzyme preference is more akin to that of other members of the ferredoxin reductase family.

6.1.3. Synthetic Chemistry with P450s

A potential application for fusion P450s is in industrial applications. Enzymes have, for many years, been commonly used in a number of industries including food production[213]. P450 enzymes with their powerful oxidation reactions and high regio- and stereo-selectivities, are potentially of use in the industrial synthesis of chemical products. Enzymatic reactions can produce chemical products of higher purity, for a lower energy input, than traditional chemical synthetic processes. In these cases enzymatic reactions could eliminate the need for conditions of high temperature, pressure or pH and the use of organic solvents. This can potentially make the synthetic process cheaper, safer and less damaging to the environment[214]. The difficulty with the use of enzymes in such applications is that they must be robust and fit into the industrial process.

As detailed in previous sections of this thesis, efforts at engineering the fusion P450 enzyme BM3 (CYP102A1) have succeeded in altering the substrate specificity to include a wide range of organic compounds. Much of the successful work in this area could also be applied to CYP102A3. Additionally, chimeric enzymes have been created based on recombination of the amino acid sequences of multiple enzymes of the CYP102 family[97]. This work explored the potential to 'mix and match' elements of the different enzymes to find an optimal combination, in this case for thermostability. Further work on CYP102A3 should be in parallel with work on BM3 (and other CYP102 enzymes) and aimed at identifying desirable aspects of each enzyme and ultimately developing new applications for them.

6.1.4. Bioremediation Using P450s

The most immediate potential application for CYP116B1, is in the field of bioremediation. This describes the use of biological systems in the removal of dangerous or undesirable substances from the environment. The use of P450 enzymes is in keeping with their physiological roles, frequently performing steps in detoxification pathways. In general, applications of this type are effected by engineering the gene coding for an appropriate enzyme into a host organism (bacterial, fungal or plant) which is then introduced into the contaminated

environment. It is desirable, if the inserted gene is to be retained by the host organism, that the gene conveys some survival advantage upon it.

The use of fusion P450s for this kind of application has already been demonstrated in principle using CYP177A1 (XplA), a P450-flavodoxin fusion with the ability to degrade the explosive compound RDX. Plants of the *Arabidopsis* genus engineered to express the *xplA* gene were shown to have the ability to remove RDX from the soil in which they were grown[204]. In this case the ability to metabolise RDX is also beneficial to the host species as the explosive is used as a nitrogen source, and the plants grown in RDX contaminated soil showed better growth than control groups.

In this model of a bioremediation system the advantages of using a fusion P450 enzyme are clear. Fusion P450s represent a single gene to be engineered into a host organism as opposed to multiple genes for P450 enzyme and its redox partner(s). It is possible, in many cases, that multiple transgenic enzymes (of which the P450 is one) will be required to allow full mineralisation of the target compound(s) by the host organism. Additionally, as demonstrated in the model fusion P450, BM3, the fused domain arrangement allows more efficient electron transfer between the haem and reductase domains compared to P450s with separate redox partners[71]. Overall, this represents a more efficient, faster reaction which could translate to more effective bioremediation.

Further work would be required to produce a similarly successful bioremediation system using CYP116B1. The enzyme already has demonstrable activity against common soil contaminants: the two thiocarbamate pesticides EPTC and vernolate. Efforts with this enzyme might focus on widening the range of substrates, starting with other thiocarbamates which at present do not appear to be oxidised by the enzyme, and moving on to less closely related substrates. In the future, further work on this family of enzymes could see the creation of P450 fusion enzymes targeting a wide range of environmental contaminants as substrates.

6.1.5. Suggestions for Future Studies

Further work on the CYP102A3 enzyme should include further efforts to produce a structure of the isolated haem domain. This would be best attempted by seeking

conditions under which an alternative crystal form, which could be used successfully for x-ray diffraction experiments, is produced. A haem domain structure would show what structural differences exist between CYP102A3 and BM3, two enzymes with high amino acid sequence identity. The structure would also pave the way for many other studies on this enzyme that could provide further explanation for the variations in substrate specificity, stereo- and regio-selectivity between enzymes of the CYP102A subfamily. These studies could include mutagenesis studies and *in silico* modelling. It is clear that further examination of the key differences in amino acid sequence between CYP102A3 and BM3 could assist attempts to engineer the reactivity of both these enzymes and cytochromes P450 in general.

In the case of CYP116B1, further examination of the unusual imidazole binding and the relatively tight binding of 2-phenylimidazole to the haem domain might provide information into the nature of the substrate binding cavity of this enzyme. This could be achieved through structural characterisation of ligand-bound species of the isolated haem domain, and also by using isothermal titration calorimetry to explore ligand affinity and stoichiometry of binding. This last technique might also be used to probe the true extent of substrate binding and substrate approach to the distal haem ligand, with reference to the small effect upon the haem iron spin-state and redox potentials, identified in this thesis, on binding the EPTC and vernolate substrates.

Further members of the CYP116B1 substrate library, including the substrate candidates looked at in this work, could be used to further investigate why many substrate candidates appear to stimulate NADPH oxidation without undergoing oxidation by the enzyme, which could reflect the absence of binding to CYP116B1 or a binding mode distant from the haem-iron.

With regards to the reductase domain of CYP116B1, the first step should be to develop a more reliable purification scheme, possibly by employing the same 6xHis tag system that was used for the FLCYP116B1 and HDCYP116B1 proteins. Structural characterisation of the RDCYP116B1 protein could further explore the homology with PDOR and the wider ferredoxin reductase family. Pre-steady state kinetic experiments could also be conducted to explore the individual steps of the reductase reaction. By reference to other work on FNR enzymes, this could provide more information about the evolution of this class of redox enzymes as a whole.

The work at the beginning of chapter 4 identifies a whole range of other putative P450 fusion enzymes that may warrant further study. Investigation of these may reveal further novel redox partner systems for P450 enzymes which could be interesting and potentially useful for the two applications described above, amongst others. Research on P450 fusion enzymes has so far provided a great deal of information about the cytochromes P450 in general. As our understanding of these enzymes is both biotechnologically and medically relevant, the study of P450 fusions should definitely be pursued in future research.

6.2. Bibliography

1. Lewis, C.A., Jr. and R. Wolfenden, *Uroporphyrinogen decarboxylation as a benchmark for the catalytic proficiency of enzymes*. Proc Natl Acad Sci U S A, 2008. **105**(45): p. 17328-33.
2. Stryer, L., *Biochemistry*. 4th ed. ed. 1995, New York: W.H. Freeman. xxxiv,1064p.
3. Mansoorabadi, S.O., C.J. Thibodeaux, and H.W. Liu, *The diverse roles of flavin coenzymes--nature's most versatile thespians*. J Org Chem, 2007. **72**(17): p. 6329-42.
4. Hirst, J., *Towards the molecular mechanism of respiratory complex I*. Biochem J, 2010. **425**(2): p. 327-39.
5. Phillips, I.R. and E.A. Shephard, *Flavin-containing monooxygenases: mutations, disease and drug response*. Trends Pharmacol Sci, 2008. **29**(6): p. 294-301.
6. Chapman, S.K. and G.A. Reid, *Flavoprotein protocols*. 1999, Totowa, N.J.: Humana. xii, 256 p.
7. Dubois, J.L. and J.P. Klinman, *Mechanism of post-translational quinone formation in copper amine oxidases and its relationship to the catalytic turnover*. Arch Biochem Biophys, 2005. **433**(1): p. 255-65.
8. Gunner, M.R., J. Madeo, and Z. Zhu, *Modification of quinone electrochemistry by the proteins in the biological electron transfer chains: examples from photosynthetic reaction centers*. J Bioenerg Biomembr, 2008. **40**(5): p. 509-19.
9. Ferraro, D.J., L. Gakhar, and S. Ramaswamy, *Rieske business: structure-function of Rieske non-heme oxygenases*. Biochem Biophys Res Commun, 2005. **338**(1): p. 175-90.
10. Solomon, E.I., X. Xie, and A. Dey, *Mixed valent sites in biological electron transfer*. Chem Soc Rev, 2008. **37**(4): p. 623-38.
11. Arnon, D.I., *The discovery of ferredoxin: the photosynthetic path*. Trends Biochem Sci, 1988. **13**(1): p. 30-3.
12. Vassiliev, I.R., M.L. Antonkine, and J.H. Golbeck, *Iron-sulfur clusters in type I reaction centers*. Biochim Biophys Acta, 2001. **1507**(1-3): p. 139-60.
13. Guallar, V. and B. Olsen, *The role of the heme propionates in heme biochemistry*. J Inorg Biochem, 2006. **100**(4): p. 755-60.
14. Caughey, W.S., G.A. Smythe, D.H. O'Keeffe, J.E. Maskasky, and M.I. Smith, *Heme A of cytochrome c oxidase. Structure and properties: comparisons with hemes B, C, and S and derivatives*. Journal of Biological Chemistry, 1975. **250**(19): p. 7602-7622.
15. McLean, M.A., S.A. Maves, K.E. Weiss, S. Krepich, and S.G. Sligar, *Characterization of a cytochrome P450 from the acidothermophilic archaea Sulfolobus solfataricus*. Biochem Biophys Res Commun, 1998. **252**(1): p. 166-72.
16. Omura, T. and R. Sato, *The Carbon Monoxide-Binding Pigment of Liver Microsomes. II. Solubilization, Purification, and Properties*. J Biol Chem, 1964. **239**: p. 2379-85.
17. Omura, T. and R. Sato, *The Carbon Monoxide-Binding Pigment of Liver Microsomes. I. Evidence for Its Hemoprotein Nature*. J Biol Chem, 1964. **239**: p. 2370-8.
18. Nebert, D.W., M. Adesnik, M.J. Coon, R.W. Estabrook, F.J. Gonzalez, F.P. Guengerich, I.C. Gunsalus, E.F. Johnson, B. Kemper, W. Levin, and et al., *The P450 gene superfamily: recommended nomenclature*. DNA, 1987. **6**(1): p. 1-11.

19. Nelson, D.R., T. Kamataki, D.J. Waxman, F.P. Guengerich, R.W. Estabrook, R. Feyereisen, F.J. Gonzalez, M.J. Coon, I.C. Gunsalus, O. Gotoh, and et al., *The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature*. DNA Cell Biol, 1993. **12**(1): p. 1-51.
20. Nelson, D.R., *Cytochrome P450 nomenclature*. Methods Mol Biol, 1998. **107**: p. 15-24.
21. Nelson, D.R., *Cytochrome P450 nomenclature, 2004*. Methods Mol Biol, 2006. **320**: p. 1-10.
22. Nelson, D.R., *The cytochrome p450 homepage*. Hum Genomics, 2009. **4**(1): p. 59-65.
23. Guengerich, F.P., Z.L. Wu, and C.J. Bartleson, *Function of human cytochrome P450s: characterization of the orphans*. Biochem Biophys Res Commun, 2005. **338**(1): p. 465-9.
24. Iyanagi, T., *Molecular mechanism of phase I and phase II drug-metabolizing enzymes: implications for detoxification*. Int Rev Cytol, 2007. **260**: p. 35-112.
25. Lee, S.S., J.T. Buters, T. Pineau, P. Fernandez-Salguero, and F.J. Gonzalez, *Role of CYP2E1 in the hepatotoxicity of acetaminophen*. J Biol Chem, 1996. **271**(20): p. 12063-7.
26. Prescott, L.F., *Paracetamol, alcohol and the liver*. Br J Clin Pharmacol, 2000. **49**(4): p. 291-301.
27. Nebert, D.W. and D.W. Russell, *Clinical importance of the cytochromes P450*. Lancet, 2002. **360**(9340): p. 1155-62.
28. Subramanian, A., M. Salhab, and K. Mokbel, *Oestrogen producing enzymes and mammary carcinogenesis: a review*. Breast Cancer Res Treat, 2008. **111**(2): p. 191-202.
29. Lepesheva, G.I. and M.R. Waterman, *Sterol 14alpha-demethylase cytochrome P450 (CYP51), a P450 in all biological kingdoms*. Biochim Biophys Acta, 2007. **1770**(3): p. 467-77.
30. Hughes, I.A., *Congenital adrenal hyperplasia*, in *Oxford textbook of medicine*, D.A. Warrell, T.M. Cox, and J.D. Firth, Editors. 2005, Oxford University Press, 2003 (2005 printing): Oxford. p. 261-264.
31. Peter Guengerich, F., *Cytochrome P450: What Have We Learned and What Are the Future Issues?* Drug Metabolism Reviews, 2004. **36**(2): p. 159-197.
32. Lamb, D.C., M.R. Waterman, S.L. Kelly, and F.P. Guengerich, *Cytochromes P450 and drug discovery*. Current Opinion in Biotechnology, 2007. **18**(6): p. 504-512.
33. *TubercuList, Genome of Mycobacterium tuberculosis*. 2008 01/10/2008 [cited 2010 14/04/2010]; A search for partial gene name using the keyword 'cyp']. Available from: <http://genolist.pasteur.fr/TubercuList/>.
34. McLean, K.J., D. Clift, D.G. Lewis, M. Sabri, P.R. Balding, M.J. Sutcliffe, D. Leys, and A.W. Munro, *The preponderance of P450s in the Mycobacterium tuberculosis genome*. Trends Microbiol, 2006. **14**(5): p. 220-8.
35. Ouellet, H., J.B. Johnston, and P.R. Ortiz de Montellano, *The Mycobacterium tuberculosis cytochrome P450 system*. Arch Biochem Biophys, 2010. **493**(1): p. 82-95.
36. Kelly, S.L., D.C. Lamb, M. Cannieux, D. Greetham, C.J. Jackson, T. Marczylo, C. Ugochukwu, and D.E. Kelly, *An old activity in the cytochrome P450 superfamily (CYP51) and a new story of drugs and resistance*. Biochem Soc Trans, 2001. **29**(Pt 2): p. 122-8.
37. McLean, K.J., A.J. Dunford, M. Sabri, R. Neeli, H.M. Girvan, P.R. Balding, D. Leys, H.E. Seward, K.R. Marshall, and A.W. Munro, *CYP121, CYP51 and associated*

- redox systems in Mycobacterium tuberculosis: towards deconvoluting enzymology of P450 systems in a human pathogen.* Biochem Soc Trans, 2006. **34**(Pt 6): p. 1178-82.
38. McLean, K.J., P. Carroll, D.G. Lewis, A.J. Dunford, H.E. Seward, R. Neeli, M.R. Cheesman, L. Marsollier, P. Douglas, W.E. Smith, I. Rosenkrands, S.T. Cole, D. Leys, T. Parish, and A.W. Munro, *Characterization of active site structure in CYP121. A cytochrome P450 essential for viability of Mycobacterium tuberculosis H37Rv.* J Biol Chem, 2008. **283**(48): p. 33406-16.
 39. Belin, P., M.H. Le Du, A. Fielding, O. Lequin, M. Jacquet, J.B. Charbonnier, A. Lecoq, R. Thai, M. Courcon, C. Masson, C. Dugave, R. Genet, J.L. Pernodet, and M. Gondry, *Identification and structural basis of the reaction catalyzed by CYP121, an essential cytochrome P450 in Mycobacterium tuberculosis.* Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7426-31.
 40. Munro, A.W., H.M. Girvan, and K.J. McLean, *Variations on a (t)heme--novel mechanisms, redox partners and catalytic functions in the cytochrome P450 superfamily.* Nat Prod Rep, 2007. **24**(3): p. 585-609.
 41. Guengerich, F.P., *Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity.* Chem Res Toxicol, 2001. **14**(6): p. 611-50.
 42. Isin, E.M. and F.P. Guengerich, *Complex reactions catalyzed by cytochrome P450 enzymes.* Biochim Biophys Acta, 2007. **1770**(3): p. 314-29.
 43. Kumaki, K., M. Sato, H. Kon, and D.W. Nebert, *Correlation of type I, type II, and reverse type I difference spectra with absolute changes in spin state of hepatic microsomal cytochrome P-450 iron from five mammalian species.* J Biol Chem, 1978. **253**(4): p. 1048-58.
 44. Denisov, I.G., T.M. Makris, S.G. Sligar, and I. Schlichting, *Structure and chemistry of cytochrome P450.* Chem Rev, 2005. **105**(6): p. 2253-77.
 45. Gibson, G.G. and P.P. Tamburini, *Cytochrome P-450 spin state: inorganic biochemistry of haem iron ligation and functional significance.* Xenobiotica, 1984. **14**(1-2): p. 27-47.
 46. Passardi, F., C. Cosio, C. Penel, and C. Dunand, *Peroxidases have more functions than a Swiss army knife.* Plant Cell Rep, 2005. **24**(5): p. 255-65.
 47. Poulos, T.L., *Intermediates in P450 catalysis.* Philos Transact A Math Phys Eng Sci, 2005. **363**(1829): p. 793-806; discussion 1035-40.
 48. Makris, T.M., K. von Koenig, I. Schlichting, and S.G. Sligar, *The status of high-valent metal oxo complexes in the P450 cytochromes.* J Inorg Biochem, 2006. **100**(4): p. 507-18.
 49. Schlichting, I., J. Berendzen, K. Chu, A.M. Stock, S.A. Maves, D.E. Benson, R.M. Sweet, D. Ringe, G.A. Petsko, and S.G. Sligar, *The catalytic pathway of cytochrome p450cam at atomic resolution.* Science, 2000. **287**(5458): p. 1615-22.
 50. Groves, J.T. and G.A. McClusky, *Aliphatic hydroxylation by highly purified liver microsomal cytochrome P-450. Evidence for a carbon radical intermediate.* Biochem Biophys Res Commun, 1978. **81**(1): p. 154-60.
 51. Groves, J.T. and D.V. Adhyam, *Hydroxylation by cytochrome P-450 and metalloporphyrin models. Evidence for allylic rearrangement.* Journal of the American Chemical Society, 1984. **106**(7): p. 2177-2181.
 52. Hersleth, H.P., U. Ryde, P. Rydberg, C.H. Gorbitz, and K.K. Andersson, *Structures of the high-valent metal-ion haem-oxygen intermediates in peroxidases, oxygenases and catalases.* J Inorg Biochem, 2006. **100**(4): p. 460-76.

53. Ortiz de Montellano, P.R., *Cytochrome P450 : structure, mechanism, and biochemistry*. 3rd ed. ed. 2005, New York: Kluwer Academic/Plenum Publishers.
54. Hedegaard, J. and I.C. Gunsalus, *Mixed Function Oxidation*. Journal of Biological Chemistry, 1965. **240**(10): p. 4038-4043.
55. Dus, K., M. Katagiri, C.A. Yu, D.L. Erbes, and I.C. Gunsalus, *Chemical characterization of cytochrome P-450cam*. Biochemical and Biophysical Research Communications, 1970. **40**(6): p. 1423-1430.
56. Poulos, T.L., B.C. Finzel, I.C. Gunsalus, G.C. Wagner, and J. Kraut, *The 2.6-Å crystal structure of Pseudomonas putida cytochrome P-450*. J Biol Chem, 1985. **260**(30): p. 16122-30.
57. Poulos, T.L., B.C. Finzel, and A.J. Howard, *High-resolution crystal structure of cytochrome P450cam*. J Mol Biol, 1987. **195**(3): p. 687-700.
58. Podust, L.M., L.V. Yermalitskaya, G.I. Lepesheva, V.N. Podust, E.A. Dalmasso, and M.R. Waterman, *Estradiol bound and ligand-free structures of sterol 14 α -demethylase*. Structure, 2004. **12**(11): p. 1937-45.
59. Yano, J.K., L.S. Koo, D.J. Schuller, H. Li, P.R. Ortiz de Montellano, and T.L. Poulos, *Crystal structure of a thermophilic cytochrome P450 from the archaeon Sulfolobus solfataricus*. J Biol Chem, 2000. **275**(40): p. 31086-92.
60. Li, L., Z. Chang, Z. Pan, Z.Q. Fu, and X. Wang, *Modes of heme binding and substrate access for cytochrome P450 CYP74A revealed by crystal structures of allene oxide synthase*. Proc Natl Acad Sci U S A, 2008. **105**(37): p. 13883-8.
61. Yano, J.K., M.H. Hsu, K.J. Griffin, C.D. Stout, and E.F. Johnson, *Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen*. Nat Struct Mol Biol, 2005. **12**(9): p. 822-3.
62. McLean, K.J., M. Sabri, K.R. Marshall, R.J. Lawson, D.G. Lewis, D. Clift, P.R. Balding, A.J. Dunford, A.J. Warman, J.P. McVey, A.M. Quinn, M.J. Sutcliffe, N.S. Scrutton, and A.W. Munro, *Biodiversity of cytochrome P450 redox systems*. Biochem Soc Trans, 2005. **33**(Pt 4): p. 796-801.
63. Porter, T.D., *An unusual yet strongly conserved flavoprotein reductase in bacteria and mammals*. Trends Biochem Sci, 1991. **16**(4): p. 154-8.
64. Omura, T., E. Sanders, R.W. Estabrook, D.Y. Cooper, and O. Rosenthal, *Isolation from adrenal cortex of a nonheme iron protein and a flavoprotein functional as a reduced triphosphopyridine nucleotide-cytochrome P-450 reductase*. Archives of Biochemistry and Biophysics, 1966. **117**(3): p. 660-673.
65. Narhi, L.O. and A.J. Fulco, *Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P-450 monooxygenase induced by barbiturates in Bacillus megaterium*. J Biol Chem, 1986. **261**(16): p. 7160-9.
66. Nakahara, K., T. Tanimoto, K. Hatano, K. Usuda, and H. Shoun, *Cytochrome P-450 55A1 (P-450dNIR) acts as nitric oxide reductase employing NADH as the direct electron donor*. J Biol Chem, 1993. **268**(11): p. 8350-5.
67. Narhi, L.O. and A.J. Fulco, *Identification and characterization of two functional domains in cytochrome P-450BM-3, a catalytically self-sufficient monooxygenase induced by barbiturates in Bacillus megaterium*. J Biol Chem, 1987. **262**(14): p. 6683-90.
68. Miles, J.S., A.W. Munro, B.N. Rospendowski, W.E. Smith, J. McKnight, and A.J. Thomson, *Domains of the catalytically self-sufficient cytochrome P-450 BM-3. Genetic construction, overexpression, purification and spectroscopic characterization*. Biochem J, 1992. **288** (Pt 2): p. 503-9.

69. Noble, M.A., C.S. Miles, S.K. Chapman, D.A. Lysek, A.C. MacKay, G.A. Reid, R.P. Hanzlik, and A.W. Munro, *Roles of key active-site residues in flavocytochrome P450 BM3*. *Biochem J*, 1999. **339 (Pt 2)**: p. 371-9.
70. Schneider, S., M.G. Wubbolts, D. Sanglard, and B. Witholt, *Biocatalyst engineering by assembly of fatty acid transport and oxidation activities for In vivo application of cytochrome P-450BM-3 monooxygenase*. *Appl Environ Microbiol*, 1998. **64**(10): p. 3784-90.
71. Munro, A.W., D.G. Leys, K.J. McLean, K.R. Marshall, T.W. Ost, S. Daff, C.S. Miles, S.K. Chapman, D.A. Lysek, C.C. Moser, C.C. Page, and P.L. Dutton, *P450 BM3: the very model of a modern flavocytochrome*. *Trends Biochem Sci*, 2002. **27**(5): p. 250-7.
72. Black, S.D. and S.T. Martin, *Evidence for conformational dynamics and molecular aggregation in cytochrome P450 102 (BM-3)*. *Biochemistry*, 1994. **33**(40): p. 12056-62.
73. Neeli, R., H.M. Girvan, A. Lawrence, M.J. Warren, D. Leys, N.S. Scrutton, and A.W. Munro, *The dimeric form of flavocytochrome P450 BM3 is catalytically functional as a fatty acid hydroxylase*. *FEBS Lett*, 2005. **579**(25): p. 5582-8.
74. Kitazume, T., D.C. Haines, R.W. Estabrook, B. Chen, and J.A. Peterson, *Obligatory Intermolecular Electron-Transfer from FAD to FMN in Dimeric P450BM-3†*. *Biochemistry*, 2007. **46**(42): p. 11892-11901.
75. Nakayama, N., A. Takemae, and H. Shoun, *Cytochrome P450foxy, a catalytically self-sufficient fatty acid hydroxylase of the fungus Fusarium oxysporum*. *J Biochem*, 1996. **119**(3): p. 435-40.
76. Kitazume, T., N. Takaya, N. Nakayama, and H. Shoun, *Fusarium oxysporum fatty-acid subterminal hydroxylase (CYP505) is a membrane-bound eukaryotic counterpart of Bacillus megaterium cytochrome P450BM3*. *J Biol Chem*, 2000. **275**(50): p. 39734-40.
77. Kitazume, T., A. Tanaka, N. Takaya, A. Nakamura, S. Matsuyama, T. Suzuki, and H. Shoun, *Kinetic analysis of hydroxylation of saturated fatty acids by recombinant P450foxy produced by an Escherichia coli expression system*. *Eur J Biochem*, 2002. **269**(8): p. 2075-82.
78. Gustafsson, M.C., C.N. Palmer, C.R. Wolf, and C. von Wachenfeldt, *Fatty-acid-displaced transcriptional repressor, a conserved regulator of cytochrome P450 102 transcription in Bacillus species*. *Arch Microbiol*, 2001. **176**(6): p. 459-64.
79. Gustafsson, M.C., O. Roitel, K.R. Marshall, M.A. Noble, S.K. Chapman, A. Pessegueiro, A.J. Fulco, M.R. Cheesman, C. von Wachenfeldt, and A.W. Munro, *Expression, purification, and characterization of Bacillus subtilis cytochromes P450 CYP102A2 and CYP102A3: flavocytochrome homologues of P450 BM3 from Bacillus megaterium*. *Biochemistry*, 2004. **43**(18): p. 5474-87.
80. Lentz, O., V. Urlacher, and R.D. Schmid, *Substrate specificity of native and mutated cytochrome P450 (CYP102A3) from Bacillus subtilis*. *J Biotechnol*, 2004. **108**(1): p. 41-9.
81. Nelson, D. *Cytochrome P450 Homepage*. 2010 25/01/2010 [cited 2010 05/04/2010]; A website containing information about the cytochromes P450]. Available from: <http://drnelson.utmem.edu/CytochromeP450.html>.
82. Chenna, R., H. Sugawara, T. Koike, R. Lopez, T.J. Gibson, D.G. Higgins, and J.D. Thompson, *Multiple sequence alignment with the Clustal series of programs*. *Nucleic Acids Res*, 2003. **31**(13): p. 3497-500.
83. Jackson, C.J., D.C. Lamb, T.H. Marczylo, A.G. Warrilow, N.J. Manning, D.J. Lowe, D.E. Kelly, and S.L. Kelly, *A novel sterol 14alpha-demethylase/ferredoxin fusion*

- protein (MCCYP51FX) from *Methylococcus capsulatus* represents a new class of the cytochrome P450 superfamily. *J Biol Chem*, 2002. **277**(49): p. 46959-65.
84. Munro, A.W., H.M. Girvan, and K.J. McLean, *Cytochrome P450--redox partner fusion enzymes*. *Biochim Biophys Acta*, 2007. **1770**(3): p. 345-59.
 85. Ravichandran, K.G., S.S. Boddupalli, C.A. Hasermann, J.A. Peterson, and J. Deisenhofer, *Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's*. *Science*, 1993. **261**(5122): p. 731-6.
 86. Munro, A.W., K. Malarkey, J. McKnight, A.J. Thomson, S.M. Kelly, N.C. Price, J.G. Lindsay, J.R. Coggins, and J.S. Miles, *The role of tryptophan 97 of cytochrome P450 BM3 from *Bacillus megaterium* in catalytic function. Evidence against the 'covalent switching' hypothesis of P-450 electron transfer*. *Biochem J*, 1994. **303** (Pt 2): p. 423-8.
 87. Li, H. and T.L. Poulos, *The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid*. *Nat Struct Biol*, 1997. **4**(2): p. 140-6.
 88. Oliver, C.F., S. Modi, M.J. Sutcliffe, W.U. Primrose, L.Y. Lian, and G.C. Roberts, *A single mutation in cytochrome P450 BM3 changes substrate orientation in a catalytic intermediate and the regioselectivity of hydroxylation*. *Biochemistry*, 1997. **36**(7): p. 1567-72.
 89. Dietrich, M., T.A. Do, R.D. Schmid, J. Pleiss, and V.B. Urlacher, *Altering the regioselectivity of the subterminal fatty acid hydroxylase P450 BM-3 towards [gamma]- and [delta]-positions*. *Journal of Biotechnology*, 2009. **139**(1): p. 115-117.
 90. Ost, T.W., C.S. Miles, J. Murdoch, Y. Cheung, G.A. Reid, S.K. Chapman, and A.W. Munro, *Rational re-design of the substrate binding site of flavocytochrome P450 BM3*. *FEBS Lett*, 2000. **486**(2): p. 173-7.
 91. Carmichael, A.B. and L.L. Wong, *Protein engineering of *Bacillus megaterium* CYP102*. *European Journal of Biochemistry*, 2001. **268**(10): p. 3117-3125.
 92. Meinhold, P., M.W. Peters, M.M.Y. Chen, K. Takahashi, and F.H. Arnold, *Direct Conversion of Ethane to Ethanol by Engineered Cytochrome P450 BM3*. *Chembiochem*, 2005. **6**(10): p. 1765-1768.
 93. Fasan, R., M.M. Chen, N.C. Crook, and F.H. Arnold, *Engineered alkane-hydroxylating cytochrome P450(BM3) exhibiting natively like catalytic properties*. *Angew Chem Int Ed Engl*, 2007. **46**(44): p. 8414-8.
 94. Whitehouse, C.J.C., S.G. Bell, H.G. Tufton, R.J.P. Kenny, L.C.I. Ogilvie, and L.-L. Wong, *Evolved CYP102A1 (P450BM3) variants oxidise a range of non-natural substrates and offer new selectivity options*. *Chemical Communications*, 2008(8): p. 966-968.
 95. Sawayama, A.M., M.M. Chen, P. Kulanthaivel, M.S. Kuo, H. Hemmerle, and F.H. Arnold, *A panel of cytochrome P450 BM3 variants to produce drug metabolites and diversify lead compounds*. *Chemistry*, 2009. **15**(43): p. 11723-9.
 96. Otey, C.R., M. Landwehr, J.B. Endelman, K. Hiraga, J.D. Bloom, and F.H. Arnold, *Structure-guided recombination creates an artificial family of cytochromes P450*. *PLoS Biol*, 2006. **4**(5): p. e112.
 97. Eiben, S., H. Bartelmas, and V.B. Urlacher, *Construction of a thermostable cytochrome P450 chimera derived from self-sufficient mesophilic parents*. *Appl Microbiol Biotechnol*, 2007. **75**(5): p. 1055-61.
 98. Dailey, H.A., T.A. Dailey, C.K. Wu, A.E. Medlock, K.F. Wang, J.P. Rose, and B.C. Wang, *Ferrocyclase at the millennium: structures, mechanisms and [2Fe-2S] clusters*. *Cell Mol Life Sci*, 2000. **57**(13-14): p. 1909-26.

99. Michaelis, L. and M.L. Menten, *Die kinetik der invertinwirkung*. Biochem. Z, 1913. **49**(333-369): p. 352.
100. Atkins, P.W. and J. De Paula, *Atkins' physical chemistry*. 8th ed. ed. 2006, Oxford: Oxford University Press. xxx, 1064 p.
101. Robert, B., *Resonance Raman spectroscopy*. Photosynth Res, 2009. **101**(2-3): p. 147-55.
102. Desbois, A., G. Mazza, F. Stetzkowski, and M. Lutz, *Resonance raman spectroscopy of protoheme-protein interactions in oxygen-carrying hemoproteins and in peroxidases*. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 1984. **785**(3): p. 161-176.
103. Desbois, A., *Resonance Raman spectroscopy of c-type cytochromes*. Biochimie, 1994. **76**(7): p. 693-707.
104. Hudecek, J., E. Anzenbacherova, P. Anzenbacher, A.W. Munro, and P. Hildebrandt, *Structural similarities and differences of the heme pockets of various P450 isoforms as revealed by resonance Raman spectroscopy*. Arch Biochem Biophys, 2000. **383**(1): p. 70-8.
105. Spiro, T.G. and T.C. Streckas, *Resonance Raman spectra of heme proteins. Effects of oxidation and spin state*. J Am Chem Soc, 1974. **96**(2): p. 338-45.
106. Girvan, H.M., C.W. Levy, P. Williams, K. Fisher, M.R. Cheesman, S.E. Rigby, D. Leys, and A.W. Munro, *Glutamate-haem ester bond formation is disfavoured in flavocytochrome P450 BM3: characterization of glutamate substitution mutants at the haem site of P450 BM3*. Biochem J, 2010. **427**(3): p. 455-66.
107. Weil, J.A. and J.R. Bolton, *Electron paramagnetic resonance : elementary theory and practical applications*. 2nd ed. / John A. Weil, James R. Bolton. ed. 2007, Hoboken, N.J.: Wiley ; Chichester : John Wiley [distributor]. xxiii, 664 p.
108. Dutton, P.L., *Redox potentiometry: determination of midpoint potentials of oxidation-reduction components of biological electron-transfer systems*. Methods Enzymol, 1978. **54**: p. 411-35.
109. Mayhew, S.G., *Potentiometric Measurement of Oxidation-Reduction Potentials*, in *Flavoprotein Protocols*, S.K. Chapman and G.A. Reid, Editors. 1999, Humana Press. p. 49-59.
110. Gunsalus, I.C., T.C. Pederson, and S.G. Sligar, *Oxygenase-catalyzed biological hydroxylations*. Annu Rev Biochem, 1975. **44**: p. 377-407.
111. Daff, S.N., S.K. Chapman, K.L. Turner, R.A. Holt, S. Govindaraj, T.L. Poulos, and A.W. Munro, *Redox control of the catalytic cycle of flavocytochrome P-450 BM3*. Biochemistry, 1997. **36**(45): p. 13816-23.
112. Munro, A.W., J.G. Lindsay, J.R. Coggins, S.M. Kelly, and N.C. Price, *Analysis of the structural stability of the multidomain enzyme flavocytochrome P-450 BM3*. Biochim Biophys Acta, 1996. **1296**(2): p. 127-37.
113. Kunst, F., N. Ogasawara, I. Moszer, A.M. Albertini, G. Alloni, V. Azevedo, M.G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S.C. Brignell, S. Bron, S. Brouillet, C.V. Bruschi, B. Caldwell, V. Capuano, N.M. Carter, S.K. Choi, J.J. Codani, I.F. Connerton, A. Danchin, and et al., *The complete genome sequence of the gram-positive bacterium Bacillus subtilis*. Nature, 1997. **390**(6657): p. 249-56.
114. Lee, T.R., H.P. Hsu, and G.C. Shaw, *Transcriptional regulation of the Bacillus subtilis bscR-CYP102A3 operon by the BscR repressor and differential induction of cytochrome CYP102A3 expression by oleic acid and palmitate*. J Biochem (Tokyo), 2001. **130**(4): p. 569-74.

115. Budde, M., M. Morr, R.D. Schmid, and V.B. Urlacher, *Selective hydroxylation of highly branched fatty acids and their derivatives by CYP102A1 from Bacillus megaterium*. *Chembiochem*, 2006. **7**(5): p. 789-94.
116. Cryle, M.J., R.D. Espinoza, S.J. Smith, N.J. Matovic, and J.J. De Voss, *Are branched chain fatty acids the natural substrates for P450(BM3)?* *Chem Commun (Camb)*, 2006(22): p. 2353-5.
117. Lentz, O., A. Feenstra, T. Habicher, B. Hauer, R.D. Schmid, and V.B. Urlacher, *Altering the regioselectivity of cytochrome P450 CYP102A3 of Bacillus subtilis by using a new versatile assay system*. *Chembiochem*, 2006. **7**(2): p. 345-50.
118. Nagy, I., F. Compennolle, K. Ghys, J. Vanderleyden, and R. De Mot, *A single cytochrome P-450 system is involved in degradation of the herbicides EPTC (S-ethyl dipropylthiocarbamate) and atrazine by Rhodococcus sp. strain NI86/21*. *Appl Environ Microbiol*, 1995. **61**(5): p. 2056-60.
119. Nagy, I., G. Schoofs, F. Compennolle, P. Proost, J. Vanderleyden, and R. de Mot, *Degradation of the thiocarbamate herbicide EPTC (S-ethyl dipropylcarbamothioate) and biosafening by Rhodococcus sp. strain NI86/21 involve an inducible cytochrome P-450 system and aldehyde dehydrogenase*. *J Bacteriol*, 1995. **177**(3): p. 676-87.
120. Shao, Z.Q. and R. Behki, *Cloning of the genes for degradation of the herbicides EPTC (S-ethyl dipropylthiocarbamate) and atrazine from Rhodococcus sp. strain TE1*. *Appl Environ Microbiol*, 1995. **61**(5): p. 2061-5.
121. von Rozycki, T. and D.H. Nies, *Cupriavidus metallidurans: evolution of a metal-resistant bacterium*. *Antonie Van Leeuwenhoek*, 2009. **96**(2): p. 115-39.
122. Mergeay, M., S. Monchy, T. Vallaey, V. Auquier, A. Benotmane, P. Bertin, S. Taghavi, J. Dunn, D. van der Lelie, and R. Wattiez, *Ralstonia metallidurans, a bacterium specifically adapted to toxic metals: towards a catalogue of metal-responsive genes*. *FEMS Microbiol Rev*, 2003. **27**(2-3): p. 385-410.
123. Monchy, S., M.A. Benotmane, P. Janssen, T. Vallaey, S. Taghavi, D. van der Lelie, and M. Mergeay, *Plasmids pMOL28 and pMOL30 of Cupriavidus metallidurans are specialized in the maximal viable response to heavy metals*. *J Bacteriol*, 2007. **189**(20): p. 7417-25.
124. Reith, F., S.L. Rogers, D.C. McPhail, and D. Webb, *Bio-mineralization of gold: biofilms on bacteriiform gold*. *Science*, 2006. **313**(5784): p. 233-6.
125. Roux, M., G. Sarret, I. Pignot-Paintrand, M. Fontecave, and J. Coves, *Mobilization of selenite by Ralstonia metallidurans CH34*. *Appl Environ Microbiol*, 2001. **67**(2): p. 769-73.
126. Taghavi, S., C. Lesaulnier, S. Monchy, R. Wattiez, M. Mergeay, and D. van der Lelie, *Lead(II) resistance in Cupriavidus metallidurans CH34: interplay between plasmid and chromosomally-located functions*. *Antonie Van Leeuwenhoek*, 2009. **96**(2): p. 171-82.
127. Jian, X., E.C. Wasinger, J.V. Lockard, L.X. Chen, and C. He, *Highly sensitive and selective gold(I) recognition by a metalloregulator in Ralstonia metallidurans*. *J Am Chem Soc*, 2009. **131**(31): p. 10869-71.
128. Janssen, P.J., R. Van Houdt, H. Moors, P. Monsieurs, N. Morin, A. Michaux, M.A. Benotmane, N. Leys, T. Vallaey, A. Lapidus, S. Monchy, C. Medigue, S. Taghavi, S. McCorkle, J. Dunn, D. van der Lelie, and M. Mergeay, *The complete genome sequence of Cupriavidus metallidurans strain CH34, a master survivalist in harsh and anthropogenic environments*. *PLoS One*, 2010. **5**(5): p. e10433.
129. Correll, C.C., C.J. Batie, D.P. Ballou, and M.L. Ludwig, *Phthalate dioxygenase reductase: a modular structure for electron transfer from pyridine nucleotides to [2Fe-2S]*. *Science*, 1992. **258**(5088): p. 1604-10.

130. Correll, C.C., C.J. Batie, D.P. Ballou, and M.L. Ludwig, *Crystallographic characterization of phthalate oxygenase reductase, an iron-sulfur flavoprotein from Pseudomonas cepacia*. J Biol Chem, 1985. **260**(27): p. 14633-5.
131. Gassner, G., L. Wang, C. Batie, and D.P. Ballou, *Reaction of phthalate dioxygenase reductase with NADH and NAD: kinetic and spectral characterization of intermediates*. Biochemistry, 1994. **33**(40): p. 12184-93.
132. Gassner, G.T., M.L. Ludwig, D.L. Gatti, C.C. Correll, and D.P. Ballou, *Structure and mechanism of the iron-sulfur flavoprotein phthalate dioxygenase reductase*. Faseb J, 1995. **9**(14): p. 1411-8.
133. Stevenson, R.C., W.R. Dunham, R.H. Sands, T.P. Singer, and H. Beinert, *Studies on the spin-spin interaction between flavin and iron-sulfur cluster in an iron-sulfur flavoprotein*. Biochim Biophys Acta, 1986. **869**(1): p. 81-8.
134. Batie, C.J., E. LaHaie, and D.P. Ballou, *Purification and characterization of phthalate oxygenase and phthalate oxygenase reductase from Pseudomonas cepacia*. J Biol Chem, 1987. **262**(4): p. 1510-8.
135. Warman, A.J., *The Molecular Enzymology of Cytochromes P450 of Biotechnological and Medical Interest*, in Faculty of Life Sciences. 2006, The University of Leicester: Leicester.
136. Zimm, B.H., *The dependence of the scattering of light on angle and concentration in linear polymer solutions*. J Phys Colloid Chem, 1948. **52**(1): p. 260-7.
137. Sevrioukova, I.F., H. Li, H. Zhang, J.A. Peterson, and T.L. Poulos, *Structure of a cytochrome P450-redox partner electron-transfer complex*. Proc Natl Acad Sci U S A, 1999. **96**(5): p. 1863-8.
138. Ost, T.W., A.W. Munro, C.G. Mowat, P.R. Taylor, A. Pesseguiro, A.J. Fulco, A.K. Cho, M.A. Cheesman, M.D. Walkinshaw, and S.K. Chapman, *Structural and spectroscopic analysis of the F393H mutant of flavocytochrome P450 BM3*. Biochemistry, 2001. **40**(45): p. 13430-8.
139. Ost, T.W., J. Clark, C.G. Mowat, C.S. Miles, M.D. Walkinshaw, G.A. Reid, S.K. Chapman, and S. Daff, *Oxygen activation and electron transfer in flavocytochrome P450 BM3*. J Am Chem Soc, 2003. **125**(49): p. 15010-20.
140. Joyce, M.G., H.M. Girvan, A.W. Munro, and D. Leys, *A single mutation in cytochrome P450 BM3 induces the conformational rearrangement seen upon substrate binding in the wild-type enzyme*. J Biol Chem, 2004. **279**(22): p. 23287-93.
141. Girvan, H.M., H.E. Seward, H.S. Toogood, M.R. Cheesman, D. Leys, and A.W. Munro, *Structural and spectroscopic characterization of P450 BM3 mutants with unprecedented P450 heme iron ligand sets. New heme ligation states influence conformational equilibria in P450 BM3*. J Biol Chem, 2007. **282**(1): p. 564-72.
142. Haines, D.C., B. Chen, D.R. Tomchick, M. Bondlela, A. Hegde, M. Machius, and J.A. Peterson, *Crystal structure of inhibitor-bound P450BM-3 reveals open conformation of substrate access channel*. Biochemistry, 2008. **47**(12): p. 3662-70.
143. Girvan, H.M., H.S. Toogood, R.E. Littleford, H.E. Seward, W.E. Smith, I.S. Ekanem, D. Leys, M.R. Cheesman, and A.W. Munro, *Novel haem co-ordination variants of flavocytochrome P450BM3*. Biochem J, 2009. **417**(1): p. 65-76.
144. Whitehouse, C.J., S.G. Bell, W. Yang, J.A. Yorke, C.F. Blanford, A.J. Strong, E.J. Morse, M. Bartlam, Z. Rao, and L.L. Wong, *A highly active single-mutation variant of P450BM3 (CYP102A1)*. Chembiochem, 2009. **10**(10): p. 1654-6.

145. Chowdhary, P.K., M. Alemseghed, and D.C. Haines, *Cloning, expression and characterization of a fast self-sufficient P450: CYP102A5 from Bacillus cereus*. Arch Biochem Biophys, 2007. **468**(1): p. 32-43.
146. Dietrich, M., S. Eiben, C. Asta, T.A. Do, J. Pleiss, and V.B. Urlacher, *Cloning, expression and characterisation of CYP102A7, a self-sufficient P450 monooxygenase from Bacillus licheniformis*. Appl Microbiol Biotechnol, 2008. **79**(6): p. 931-40.
147. Lamb, D.C., L. Lei, B. Zhao, H. Yuan, C.J. Jackson, A.G. Warrilow, T. Skaug, P.J. Dyson, E.S. Dawson, S.L. Kelly, D.L. Hachey, and M.R. Waterman, *Streptomyces coelicolor A3(2) CYP102 protein, a novel fatty acid hydroxylase encoded as a heme domain without an N-terminal redox partner*. Appl Environ Microbiol, 2010. **76**(6): p. 1975-80.
148. Haines, D.C., D.R. Tomchick, M. Machius, and J.A. Peterson, *Pivotal role of water in the mechanism of P450BM-3*. Biochemistry, 2001. **40**(45): p. 13456-65.
149. Li, Q.S., J. Ogawa, R.D. Schmid, and S. Shimizu, *Engineering cytochrome P450 BM-3 for oxidation of polycyclic aromatic hydrocarbons*. Appl Environ Microbiol, 2001. **67**(12): p. 5735-9.
150. Perera, R., M. Sono, J.A. Sigman, T.D. Pfister, Y. Lu, and J.H. Dawson, *Neutral thiol as a proximal ligand to ferrous heme iron: implications for heme proteins that lose cysteine thiolate ligation on reduction*. Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3641-6.
151. Girvan, H.M., K.R. Marshall, R.J. Lawson, D. Leys, M.G. Joyce, J. Clarkson, W.E. Smith, M.R. Cheesman, and A.W. Munro, *Flavocytochrome P450 BM3 mutant A264E undergoes substrate-dependent formation of a novel heme iron ligand set*. J Biol Chem, 2004. **279**(22): p. 23274-86.
152. Ouellet, H., L.M. Podust, and P.R. de Montellano, *Mycobacterium tuberculosis CYP130: crystal structure, biophysical characterization, and interactions with antifungal azole drugs*. J Biol Chem, 2008. **283**(8): p. 5069-80.
153. Cupp-Vickery, J., R. Anderson, and Z. Hatziris, *Crystal structures of ligand complexes of P450eryF exhibiting homotropic cooperativity*. Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3050-5.
154. Ekroos, M. and T. Sjogren, *Structural basis for ligand promiscuity in cytochrome P450 3A4*. Proc Natl Acad Sci U S A, 2006. **103**(37): p. 13682-7.
155. Zhao, B., F.P. Guengerich, M. Voehler, and M.R. Waterman, *Role of active site water molecules and substrate hydroxyl groups in oxygen activation by cytochrome P450 158A2: a new mechanism of proton transfer*. J Biol Chem, 2005. **280**(51): p. 42188-97.
156. Dawson, J.H., L.A. Andersson, and M. Sono, *Spectroscopic investigations of ferric cytochrome P-450-CAM ligand complexes. Identification of the ligand trans to cysteinate in the native enzyme*. J Biol Chem, 1982. **257**(7): p. 3606-17.
157. Poulos, T.L. and A.J. Howard, *Crystal structures of metyrapone- and phenylimidazole-inhibited complexes of cytochrome P-450cam*. Biochemistry, 1987. **26**(25): p. 8165-74.
158. McKnight, J., M.R. Cheesman, A.J. Thomson, J.S. Miles, and A.W. Munro, *Identification of charge-transfer transitions in the optical spectrum of low-spin ferric cytochrome P-450 Bacillus megaterium*. Eur J Biochem, 1993. **213**(2): p. 683-7.
159. Jovanovic, T., R. Farid, R.A. Friesner, and A.E. McDermott, *Thermal equilibrium of high- and low-spin forms of cytochrome P450 BM-3: repositioning of the substrate?* J Am Chem Soc, 2005. **127**(39): p. 13548-52.

160. Matsuura, K., S. Yoshioka, T. Tosha, H. Hori, K. Ishimori, T. Kitagawa, I. Morishima, N. Kagawa, and M.R. Waterman, *Structural diversities of active site in clinical azole-bound forms between sterol 14 α -demethylases (CYP51s) from human and Mycobacterium tuberculosis*. J Biol Chem, 2005. **280**(10): p. 9088-96.
161. Sono, M. and J.H. Dawson, *Formation of low spin complexes of ferric cytochrome P-450-CAM with anionic ligands. Spin state and ligand affinity comparison to myoglobin*. J Biol Chem, 1982. **257**(10): p. 5496-502.
162. Wells, A.V., P. Li, P.M. Champion, S.A. Martinis, and S.G. Sligar, *Resonance Raman investigations of Escherichia coli-expressed Pseudomonas putida cytochrome P450 and P420*. Biochemistry, 1992. **31**(18): p. 4384-93.
163. Smith, S.J., A.W. Munro, and W.E. Smith, *Resonance Raman scattering of cytochrome P450 BM3 and effect of imidazole inhibitors*. Biopolymers, 2003. **70**(4): p. 620-7.
164. Gunsalus, I.C. and S.G. Sligar, *Redox regulation of cytochrome P450cam mixed function oxidation by putidaredoxin and camphor ligation*. Biochimie, 1976. **58**(1-2): p. 143-7.
165. Sakurai, K., H. Shimada, T. Hayashi, and T. Tsukihara, *Substrate binding induces structural changes in cytochrome P450cam*. Acta Crystallogr Sect F Struct Biol Cryst Commun, 2009. **65**(Pt 2): p. 80-3.
166. McLean, K.J., P. Lafite, C. Levy, M.R. Cheesman, N. Mast, I.A. Pikuleva, D. Leys, and A.W. Munro, *The Structure of Mycobacterium tuberculosis CYP125: molecular basis for cholesterol binding in a P450 needed for host infection*. J Biol Chem, 2009. **284**(51): p. 35524-33.
167. Pace, C.N., *The stability of globular proteins*. CRC Crit Rev Biochem, 1975. **3**(1): p. 1-43.
168. Lawson, R.J., D. Leys, M.J. Sutcliffe, C.A. Kemp, M.R. Cheesman, S.J. Smith, J. Clarkson, W.E. Smith, I. Haq, J.B. Perkins, and A.W. Munro, *Thermodynamic and biophysical characterization of cytochrome P450 Biol from Bacillus subtilis*. Biochemistry, 2004. **43**(39): p. 12410-26.
169. Chang, J.Y., *Structural heterogeneity of 6 M GdmCl-denatured proteins: implications for the mechanism of protein folding*. Biochemistry, 2009. **48**(40): p. 9340-6.
170. Shortle, D. and C. Abeygunawardana, *NMR analysis of the residual structure in the denatured state of an unusual mutant of staphylococcal nuclease*. Structure, 1993. **1**(2): p. 121-34.
171. Kazmirski, S.L., K.B. Wong, S.M. Freund, Y.J. Tan, A.R. Fersht, and V. Daggett, *Protein folding from a highly disordered denatured state: the folding pathway of chymotrypsin inhibitor 2 at atomic resolution*. Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4349-54.
172. Tangeras, A. and T. Flatmark, *In vitro binding of protoheme IX and protoporphyrin IX to components in the matrix of rat liver mitochondria*. Biochim Biophys Acta, 1979. **588**(2): p. 201-10.
173. Andersson, M., B. Wittgren, and K.G. Wahlund, *Accuracy in multiangle light scattering measurements for molar mass and radius estimations. Model calculations and experiments*. Anal Chem, 2003. **75**(16): p. 4279-91.
174. McPherson, A., *Preparation and analysis of protein crystals*. 1989: Krieger.
175. Arnold, K., L. Bordoli, J. Kopp, and T. Schwede, *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling*. Bioinformatics, 2006. **22**(2): p. 195-201.

176. Kiefer, F., K. Arnold, M. Kunzli, L. Bordoli, and T. Schwede, *The SWISS-MODEL Repository and associated resources*. Nucleic Acids Res, 2009. **37**(Database issue): p. D387-92.
177. Peitsch, M.C., T.N. Wells, D.R. Stampf, and J.L. Sussman, *The Swiss-3DImage collection and PDB-Browser on the World-Wide Web*. Trends Biochem Sci, 1995. **20**(2): p. 82-4.
178. Marchler-Bauer, A., J.B. Anderson, F. Chitsaz, M.K. Derbyshire, C. DeWeese-Scott, J.H. Fong, L.Y. Geer, R.C. Geer, N.R. Gonzales, M. Gwadz, S. He, D.I. Hurwitz, J.D. Jackson, Z. Ke, C.J. Lanczycki, C.A. Liebert, C. Liu, F. Lu, S. Lu, G.H. Marchler, M. Mullokandov, J.S. Song, A. Tasneem, N. Thanki, R.A. Yamashita, D. Zhang, N. Zhang, and S.H. Bryant, *CDD: specific functional annotation with the Conserved Domain Database*. Nucleic Acids Res, 2009. **37**(Database issue): p. D205-10.
179. Marchler-Bauer, A. and S.H. Bryant, *CD-Search: protein domain annotations on the fly*. Nucleic Acids Res, 2004. **32**(Web Server issue): p. W327-31.
180. Seth-Smith, H.M., S.J. Rosser, A. Basran, E.R. Travis, E.R. Dabbs, S. Nicklin, and N.C. Bruce, *Cloning, sequencing, and characterization of the hexahydro-1,3,5-Trinitro-1,3,5-triazine degradation gene cluster from Rhodococcus rhodochrous*. Appl Environ Microbiol, 2002. **68**(10): p. 4764-71.
181. Karlsson, A., Z.M. Beharry, D. Matthew Eby, E.D. Coulter, E.L. Neidle, D.M. Kurtz, Jr., H. Eklund, and S. Ramaswamy, *X-ray crystal structure of benzoate 1,2-dioxygenase reductase from Acinetobacter sp. strain ADP1*. J Mol Biol, 2002. **318**(2): p. 261-72.
182. Watanabe, K., S. Narimatsu, I. Yamamoto, and H. Yoshimura, *Hepatic microsomal oxygenation of aldehydes to carboxylic acids*. Biochemical and Biophysical Research Communications, 1990. **166**(3): p. 1308-1312.
183. Pao, S.S., I.T. Paulsen, and M.H. Saier, Jr., *Major facilitator superfamily*. Microbiol Mol Biol Rev, 1998. **62**(1): p. 1-34.
184. Kureishy, N., V. Sapountzi, S. Prag, N. Anilkumar, and J.C. Adams, *Fascins, and their roles in cell structure and function*. Bioessays, 2002. **24**(4): p. 350-61.
185. Bai, C., P. Sen, K. Hofmann, L. Ma, M. Goebel, J.W. Harper, and S.J. Elledge, *SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box*. Cell, 1996. **86**(2): p. 263-74.
186. Blatch, G.L. and M. Lassle, *The tetratricopeptide repeat: a structural motif mediating protein-protein interactions*. Bioessays, 1999. **21**(11): p. 932-9.
187. Small, I.D. and N. Peeters, *The PPR motif - a TPR-related motif prevalent in plant organellar proteins*. Trends Biochem Sci, 2000. **25**(2): p. 46-7.
188. Saurin, A.J., K.L.B. Borden, M.N. Boddy, and P.S. Freemont, *Does this have a familiar RING?* Trends in Biochemical Sciences, 1996. **21**(6): p. 208-214.
189. De Mot, R. and A.H. Parret, *A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes*. Trends Microbiol, 2002. **10**(11): p. 502-8.
190. Roberts, G.A., A. Celik, D.J. Hunter, T.W. Ost, J.H. White, S.K. Chapman, N.J. Turner, and S.L. Flitsch, *A self-sufficient cytochrome p450 with a primary structural organization that includes a flavin domain and a [2Fe-2S] redox center*. J Biol Chem, 2003. **278**(49): p. 48914-20.
191. Hunter, D.J., G.A. Roberts, T.W. Ost, J.H. White, S. Muller, N.J. Turner, S.L. Flitsch, and S.K. Chapman, *Analysis of the domain properties of the novel cytochrome P450 RhF*. FEBS Lett, 2005. **579**(10): p. 2215-20.
192. Çelik, A., G.A. Roberts, J.H. White, S.K. Chapman, N.J. Turner, and S.L. Flitsch, *Probing the substrate specificity of the catalytically self-sufficient cytochrome P450 RhF from a Rhodococcus sp.* Chem Commun (Camb), 2006(43): p. 4492-4.

193. Kim, Y.H., K.H. Engesser, and S.J. Kim, *Physiological, numerical and molecular characterization of alkyl ether-utilizing rhodococci*. Environ Microbiol, 2007. **9**(6): p. 1497-510.
194. Novagen, *pET System Manual*. 2006, Novagen.
195. Siddhanta, U., A. Presta, B. Fan, D. Wolan, D.L. Rousseau, and D.J. Stuehr, *Domain swapping in inducible nitric-oxide synthase. Electron transfer occurs between flavin and heme groups located on adjacent subunits in the dimer*. J Biol Chem, 1998. **273**(30): p. 18950-8.
196. Clark, J.P., C.S. Miles, C.G. Mowat, M.D. Walkinshaw, G.A. Reid, S.N. Daff, and S.K. Chapman, *The role of Thr268 and Phe393 in cytochrome P450 BM3*. Journal of Inorganic Biochemistry, 2006. **100**(5-6): p. 1075-1090.
197. Ogura, H., C.R. Nishida, U.R. Hoch, R. Perera, J.H. Dawson, and P.R. Ortiz de Montellano, *EpoK, a Cytochrome P450 Involved in Biosynthesis of the Anticancer Agents Epothilones A and B. Substrate-Mediated Rescue of a P450 Enzyme†*. Biochemistry, 2004. **43**(46): p. 14712-14721.
198. McLean, K.J., A.J. Warman, H.E. Seward, K.R. Marshall, H.M. Girvan, M.R. Cheesman, M.R. Waterman, and A.W. Munro, *Biophysical Characterization of the Sterol Demethylase P450 from Mycobacterium tuberculosis, Its Cognate Ferredoxin, and Their Interactions†*. Biochemistry, 2006. **45**(27): p. 8427-8443.
199. Brodhun, F., C. Gobel, E. Hornung, and I. Feussner, *Identification of PpoA from Aspergillus nidulans as a fusion protein of a fatty acid heme dioxygenase/peroxidase and a cytochrome P450*. J Biol Chem, 2009. **284**(18): p. 11792-805.
200. Brodhun, F., S. Schneider, C. Gobel, E. Hornung, and I. Feussner, *PpoC from Aspergillus nidulans is a fusion protein with only one active haem*. Biochem J, 2010. **425**(3): p. 553-65.
201. Smith, A.E. and A. Fitzpatrick, *Loss of five thiolcarbamate herbicides in nonsterile soils and their stability in acidic and basic solutions*. Journal of Agricultural and Food Chemistry, 1970. **18**(4): p. 720-722.
202. Dick, W.A., R.O. Ankumah, G. McClung, and N. Abou-Assaf, *Enhanced Degradation of S-Ethyl <italic>N,N</italic>-Dipropylcarbamothioate in Soil and by an Isolated Soil Microorganism*, in *Enhanced Biodegradation of Pesticides in the Environment*. 1990, American Chemical Society. p. 98-112.
203. Jackson, R.G., E.L. Rylott, D. Fournier, J. Hawari, and N.C. Bruce, *Exploring the biochemical properties and remediation applications of the unusual explosive-degrading P450 system XplA/B*. Proc Natl Acad Sci U S A, 2007. **104**(43): p. 16822-7.
204. Rylott, E.L., R.G. Jackson, J. Edwards, G.L. Womack, H.M. Seth-Smith, D.A. Rathbone, S.E. Strand, and N.C. Bruce, *An explosive-degrading cytochrome P450 activity and its targeted application for the phytoremediation of RDX*. Nat Biotechnol, 2006. **24**(2): p. 216-9.
205. Macdonald, I.D.G., A.W. Munro, and W.E. Smith, *Fatty Acid-Induced Alteration of the Porphyrin Macrocycle of Cytochrome P450 BM3*. Biophysical journal, 1998. **74**(6): p. 3241-3249.
206. McLean, K.J., A.J. Warman, H.E. Seward, K.R. Marshall, H.M. Girvan, M.R. Cheesman, M.R. Waterman, and A.W. Munro, *Biophysical characterization of the sterol demethylase P450 from Mycobacterium tuberculosis, its cognate ferredoxin, and their interactions*. Biochemistry, 2006. **45**(27): p. 8427-43.
207. McLean, K.J., P. Carroll, D.G. Lewis, A.J. Dunford, H.E. Seward, R. Neeli, M.R. Cheesman, L. Marsollier, P. Douglas, W.E. Smith, I. Rosenkrands, S.T. Cole, D.

- Leys, T. Parish, and A.W. Munro, *Characterization of Active Site Structure in CYP121*. *Journal of Biological Chemistry*, 2008. **283**(48): p. 33406-33416.
208. Cryle, M.J., S.G. Bell, and I. Schlichting, *Structural and biochemical characterization of the cytochrome P450 CypX (CYP134A1) from Bacillus subtilis: a cyclo-L-leucyl-L-leucyl dipeptide oxidase*. *Biochemistry*, 2010. **49**(34): p. 7282-96.
209. Schuphan, I., J. Rosen, and J. Casida, *Novel activation mechanism for the promutagenic herbicide diallate*. *Science*, 1979. **205**(4410): p. 1013-1015.
210. Nisimoto, Y., *Localization of cytochrome c-binding domain on NADPH-cytochrome P-450 reductase*. *J Biol Chem*, 1986. **261**(30): p. 14232-9.
211. Murataliev, M.B., R. Feyereisen, and F.A. Walker, *Electron transfer by diflavin reductases*. *Biochim Biophys Acta*, 2004. **1698**(1): p. 1-26.
212. Bruns, C.M. and P.A. Karplus, *Refined crystal structure of spinach ferredoxin reductase at 1.7 Å resolution: oxidized, reduced and 2'-phospho-5'-AMP bound states*. *J Mol Biol*, 1995. **247**(1): p. 125-45.
213. Vroemen, A.J., *Production of industrial enzymes*. *Antonie Van Leeuwenhoek*, 1983. **49**(1): p. 90-91.
214. Wong, C.H. and G.M. Whitesides, *Enzymes in synthetic organic chemistry*. 1994, Oxford: Pergamon. xvii,370p.