Characterisation of the RDX-degrading XplA/XplB Redox System from *Rhodococcus rhodochrous*

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

2012

Soi Hang Bui
## Table of Contents

- **List of Figures** 7
- **List of Tables** 8
- **Abbreviations** 9
- **Abstract** 11
- **Declaration** 12
- **Copyright Statement** 12
- **Acknowledgements** 13

### Chapter 1: Introduction

1.1: **Cytochromes P450** 16
   - 1.1.1: Physiological Roles of P450s 18
   - 1.1.1.1: Mammalian P450s 19
   - 1.1.2: Bacterial P450s 20
   - 1.1.3: Catalytic Cycle 20
   - 1.1.2: P450 Structure 25
   - 1.1.3: Biotechnological Applications of P450s 28

1.2: **Flavoproteins** 29
   - 1.2.1: Ferredoxin-NADP+ Reductases 32
   - 1.2.1.1: Structure of Ferredoxin-NADP+ Reductases 33
   - 1.2.2: Flavodoxins 35
   - 1.2.2.1: The Flavodoxin Fold 36
   - 1.2.2.2: FMN Binding 37

1.3: **P450 Redox Systems** 38
   - 1.3.1: P450 Fusion Enzymes 41
   - 1.3.2: P450-Redox Partner Interactions 42

1.4: **The XplA/XplB Redox System** 44

1.5: **Introduction to the Project** 47

### Chapter 2: Materials & Methods

2.1: **Materials** 50

2.2: **Experimental Methods** 51
   - 2.2.1: Restriction Digestions 51
   - 2.2.2: Transformations 52
   - 2.2.3: Cloning 53
   - 2.2.4: Site-directed Mutagenesis 54
## Preliminary Pages

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.5</td>
<td>Expression Trials</td>
<td>54</td>
</tr>
<tr>
<td>2.2.6</td>
<td>Expression &amp; Purification of XplA</td>
<td>56</td>
</tr>
<tr>
<td>2.2.7</td>
<td>Expression &amp; Purification of XplA-HD</td>
<td>57</td>
</tr>
<tr>
<td>2.2.8</td>
<td>Expression &amp; Purification of XplB</td>
<td>58</td>
</tr>
<tr>
<td>2.2.9</td>
<td>Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis</td>
<td>58</td>
</tr>
<tr>
<td>2.2.10</td>
<td>Preparation of FMN-free XplA</td>
<td>60</td>
</tr>
<tr>
<td>2.2.11</td>
<td>Flavin Quantification</td>
<td>60</td>
</tr>
<tr>
<td>2.2.12</td>
<td>Flavin Reconstitution</td>
<td>61</td>
</tr>
<tr>
<td>2.2.13</td>
<td>Reconstitution of the XplA/XplB Redox System</td>
<td>61</td>
</tr>
<tr>
<td>2.2.14</td>
<td>UV/Visible Spectroscopy</td>
<td>62</td>
</tr>
<tr>
<td>2.2.15</td>
<td>Fluorescence Spectroscopy</td>
<td>64</td>
</tr>
<tr>
<td>2.2.16</td>
<td>Pyridine Hemochromogen Assay</td>
<td>64</td>
</tr>
<tr>
<td>2.2.17</td>
<td>Redox Potentiometry</td>
<td>65</td>
</tr>
<tr>
<td>2.2.18</td>
<td>Stopped-flow Spectroscopy</td>
<td>66</td>
</tr>
<tr>
<td>2.2.19</td>
<td>Electron Paramagnetic Resonance Spectroscopy</td>
<td>67</td>
</tr>
<tr>
<td>2.2.20</td>
<td>Magnetic Circular Dichroism Spectroscopy</td>
<td>68</td>
</tr>
<tr>
<td>2.2.21</td>
<td>Circular Dichroism Spectroscopy</td>
<td>68</td>
</tr>
<tr>
<td>2.2.22</td>
<td>Isothermal Titration Calorimetry</td>
<td>68</td>
</tr>
<tr>
<td>2.2.23</td>
<td>Multiangle Laser Light Scattering</td>
<td>69</td>
</tr>
<tr>
<td>2.2.24</td>
<td>Thermofluor Assay</td>
<td>69</td>
</tr>
<tr>
<td>2.2.25</td>
<td>Crystallisation Trials</td>
<td>72</td>
</tr>
<tr>
<td>2.2.26</td>
<td>XplA-HD Crystallisation &amp; Structural Analysis</td>
<td>72</td>
</tr>
</tbody>
</table>

## Chapter 3: Spectroscopic Characterisation of XplA

### 3.1: Introduction

### 3.2: Expression & Purification of XplA

### 3.3: UV/Visible Spectroscopy

#### 3.3.1: Spectroscopic Properties of XplA

#### 3.3.2: Ligand Binding Studies

#### 3.3.2.1: Type I Ligands

#### 3.3.2.2: Displacement of CO by RDX

#### 3.3.2.3: Type II Ligands

#### 3.3.2.4: Ferrous Imidazole-bound Complex

#### 3.3.2.5: Sulphur Donor Ligands

#### 3.3.2.6: Reverse Type I Ligands

#### 3.3.2.7: Determination of the Trinitrobenzene Dissociation Constant

### 3.4: Pyridine Hemochromogen Method

#### 3.4.1: Determination of the Heme Extinction Coefficient

### 3.5: Electron Paramagnetic Resonance Spectroscopy

#### 3.5.1: Identification of the Heme Ligation & Spin State

### 3.6: Magnetic Circular Dichroism Spectroscopy

#### 3.6.1: MCD Analysis of XplA

### 3.7: Redox Potentiometry

#### 3.7.1: Determination of the XplA Heme Iron Reduction Potential
3.8: Stopped-flow Spectroscopy
  3.8.1: Determination of the XplA CO Dissociation Constant 117
  3.8.2: Oxyferrous Complex Characterisation 118
  3.8.3: Kinetics of Oxyferrous Complex Formation & Collapse 120

3.9: Summary 122

Chapter 4: Structural Characterisation of XplA 127

4.1: Introduction 128

4.2: Molecular Biology of the XplA-P450 Domain 129

4.3: Expression & Purification of XplA-HD 131

4.4: Biophysical Characterisation of XplA & XplA-HD 132
  4.4.1: Thermofluor Assay 133
  4.4.1.1: XplA-HD Buffer Screen 133
  4.4.1.2: XplA-HD Additive Screen 135
  4.4.1.3: Effects on XplA-HD Thermal Stability by Ligand Binding 136
  4.4.1.4: XplA Buffer Screen 139
  4.4.2: Multiangle Laser Light Scattering 139
  4.4.2.1: MALLS Analysis of XplA 140
  4.4.2.2: MALLS Analysis of XplA-HD 141
  4.4.3: Circular Dichroism Spectroscopy 142
  4.4.3.1: CD Analysis of XplA 143

4.5: X-ray Crystallographic Studies of XplA 145
  4.5.1: Crystallisation Trials of XplA 145

4.6: X-ray Crystallographic Studies of XplA-HD 147
  4.6.1: Crystallisation Trials of XplA-HD 147
  4.6.2: Structure of Imidazole-bound XplA-HD 149
  4.6.3: Refinement of Crystallisation Conditions 153
  4.6.4: Alternative Purification Protocol for XplA-HD 156
  4.6.5: Crystallisation Trials of XplA-HD Purified Using the Alternative Protocol 159

4.7: Summary 161

Chapter 5: Expression, Purification & Characterisation of the Flavoproteins of the XplA/XplB Redox System 165

5.1: Introduction 166

5.2: Molecular Biology of XplB 167

5.3: XplB Expression Trials 168

5.4: Expression & Purification of XplB 173
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Chemical structures of heme $a$, $b$ and $c$.</td>
<td>16</td>
</tr>
<tr>
<td>1.2</td>
<td>Coordination arrangement of hexa-coordinated P450 heme in the resting and CO-bound states.</td>
<td>17</td>
</tr>
<tr>
<td>1.3</td>
<td>Schematic representation of the cytochrome P450 catalytic cycle.</td>
<td>21</td>
</tr>
<tr>
<td>1.4</td>
<td>Proposed hydrogen abstraction and oxygen rebound mechanism of the P450s.</td>
<td>23</td>
</tr>
<tr>
<td>1.5</td>
<td>Electronic configurations of the $d$-orbitals of the ferric heme iron.</td>
<td>24</td>
</tr>
<tr>
<td>1.6</td>
<td>P450 structures illustrating the common three-dimensional fold.</td>
<td>27</td>
</tr>
<tr>
<td>1.7</td>
<td>Chemical structures of riboflavin and its derivatives FMN and FAD.</td>
<td>30</td>
</tr>
<tr>
<td>1.8</td>
<td>Chemical structures of the three redox states of FMN.</td>
<td>31</td>
</tr>
<tr>
<td>1.9</td>
<td>Structures of ferredoxin reductases.</td>
<td>33</td>
</tr>
<tr>
<td>1.10</td>
<td>FAD binding in \textit{E. coli}FldR.</td>
<td>35</td>
</tr>
<tr>
<td>1.11</td>
<td>Structures of bacterial flavodoxins.</td>
<td>37</td>
</tr>
<tr>
<td>1.12</td>
<td>Superposition of the structures of apo- and holoflavodoxin from \textit{Anabaena}.</td>
<td>38</td>
</tr>
<tr>
<td>1.13</td>
<td>Schematic organisation of P450 redox systems.</td>
<td>40</td>
</tr>
<tr>
<td>1.14</td>
<td>Structures of natural and covalently linked redox partner fusion enzymes.</td>
<td>42</td>
</tr>
<tr>
<td>1.15</td>
<td>Structure of the BMP/FMN domains of BM3.</td>
<td>43</td>
</tr>
<tr>
<td>1.16</td>
<td>Chemical structures of RDX and major breakdown products.</td>
<td>46</td>
</tr>
<tr>
<td>1.17</td>
<td>Structure of the heme domain of XplA.</td>
<td>47</td>
</tr>
<tr>
<td>2.1</td>
<td>Derivation of a second order rate constant from stopped-flow kinetics.</td>
<td>67</td>
</tr>
<tr>
<td>3.1</td>
<td>Time-lapse expression of \textit{R. rhodochrous} xplA in \textit{E. coli}.</td>
<td>78</td>
</tr>
<tr>
<td>3.2</td>
<td>Ni-NTA affinity chromatography purification of XplA.</td>
<td>79</td>
</tr>
<tr>
<td>3.3</td>
<td>XplA-mediated hydroxylation of indole in \textit{E. coli}.</td>
<td>80</td>
</tr>
<tr>
<td>3.4</td>
<td>Spectroscopic characterisation of purified XplA.</td>
<td>81</td>
</tr>
<tr>
<td>3.5</td>
<td>Spectroscopic characterisation of ferric, ferrous, ferrous/CO-bound and ferric/NO-bound XplA.</td>
<td>83</td>
</tr>
<tr>
<td>3.6</td>
<td>Optical titration of XplA with RDX.</td>
<td>85</td>
</tr>
<tr>
<td>3.7</td>
<td>XplA interactions with indole.</td>
<td>86</td>
</tr>
<tr>
<td>3.8</td>
<td>RDX-dependent oxidation of CO-bound ferrous heme of XplA.</td>
<td>87</td>
</tr>
<tr>
<td>3.9</td>
<td>Optical titration of XplA with sodium cyanide.</td>
<td>90</td>
</tr>
<tr>
<td>3.10</td>
<td>Optical titrations of XplA with heterocyclic secondary amines.</td>
<td>92</td>
</tr>
<tr>
<td>3.11</td>
<td>Determination of the affinity of XplA for heterocyclic secondary amines.</td>
<td>93</td>
</tr>
<tr>
<td>3.12</td>
<td>Optical titration of XplA with 4-PI.</td>
<td>95</td>
</tr>
<tr>
<td>3.13</td>
<td>Optical titration of XplA with imidazole.</td>
<td>97</td>
</tr>
<tr>
<td>3.14</td>
<td>Optical titration of ferrous XplA with imidazole.</td>
<td>99</td>
</tr>
<tr>
<td>3.15</td>
<td>Optical titration of XplA with DTT.</td>
<td>101</td>
</tr>
<tr>
<td>3.16</td>
<td>Optical titrations of XplA with sulphur donor ligands.</td>
<td>103</td>
</tr>
<tr>
<td>3.17</td>
<td>pH-dependent binding of BME to XplA.</td>
<td>104</td>
</tr>
<tr>
<td>3.18</td>
<td>Optical titration of XplA with TNB.</td>
<td>105</td>
</tr>
<tr>
<td>3.19</td>
<td>Determination of the affinity of XplA for TNB by competitive optical titrations.</td>
<td>106</td>
</tr>
<tr>
<td>3.20</td>
<td>Pyridine hemochromogen complex of XplA.</td>
<td>108</td>
</tr>
<tr>
<td>3.21</td>
<td>EPR spectra of native and ligand-bound XplA.</td>
<td>109</td>
</tr>
<tr>
<td>3.22</td>
<td>EPR spectra of RDX- and DTT-bound XplA.</td>
<td>111</td>
</tr>
<tr>
<td>3.23</td>
<td>MCD studies of XplA.</td>
<td>113</td>
</tr>
<tr>
<td>3.24</td>
<td>Spectroelectrochemical redox titration of ligand-free XplA.</td>
<td>116</td>
</tr>
<tr>
<td>3.25</td>
<td>Kinetics of the formation of the XplA Fe(II)CO complex.</td>
<td>118</td>
</tr>
<tr>
<td>3.26</td>
<td>Characterisation of the XplA oxyferrous complex.</td>
<td>120</td>
</tr>
<tr>
<td>3.27</td>
<td>Kinetics of XplA oxyferrous complex formation and collapse.</td>
<td>121</td>
</tr>
<tr>
<td>4.1</td>
<td>\textit{NdeI} restriction digestion of mutagenised pET15b(\textit{xplA}).</td>
<td>130</td>
</tr>
<tr>
<td>4.2</td>
<td>\textit{NdeI} and \textit{BanHI} double restriction digestion of pET15b(\textit{xplA-HD}).</td>
<td>130</td>
</tr>
<tr>
<td>4.3</td>
<td>Ni-NTA affinity chromatography purification of XplA-HD.</td>
<td>131</td>
</tr>
<tr>
<td>4.4</td>
<td>Spectroscopic characterisation of purified XplA-HD.</td>
<td>132</td>
</tr>
<tr>
<td>4.5</td>
<td>Thermofluor assay of XplA-HD buffer screen.</td>
<td>134</td>
</tr>
<tr>
<td>4.6</td>
<td>Thermofluor assay of XplA-HD additive screen.</td>
<td>136</td>
</tr>
<tr>
<td>4.7</td>
<td>Thermofluor assay of ligand-bound XplA-HD additive screen.</td>
<td>138</td>
</tr>
<tr>
<td>4.8</td>
<td>MALLS analysis of XplA.</td>
<td>141</td>
</tr>
<tr>
<td>4.9</td>
<td>MALLS analysis of XplA-HD.</td>
<td>142</td>
</tr>
<tr>
<td>4.10</td>
<td>CD spectra of XplA.</td>
<td>144</td>
</tr>
</tbody>
</table>
Figure 4.11: Crystallogenesis of XplA.

Figure 4.12: Formation of diffraction quality XplA crystals.

Figure 4.13: Crystallogenesis of ligand-bound XplA-HD.

Figure 4.14: Formation of diffraction quality XplA-HD crystals.

Figure 4.15: Structure of the imidazole-coordinated XplA heme domain.

Figure 4.16: Stereo view of the ligand-protein interactions observed in the XplA-HD ternary complex.

Figure 4.17: Comparison of inhibitor binding in CYP121 and XplA-HD.

Figure 4.18: Optical titrations of XplA-HD with fatty acids.

Figure 4.19: XplA-HD interactions with polyethylene glycols.

Figure 4.20: XplA-HD and histidine interactions.

Figure 4.21: Hydroxyapatite and size-exclusion chromatography purification of XplA-HD.

Figure 4.22: EPR spectrum of hydroxyapatite and size-exclusion chromatography-purified XplA-HD.

Figure 4.23: Co-crystallisation of RDX-bound XplA-HD.

Figure 5.1: Cloning of pCold TF(xplB).

Figure 5.2: Time-lapse production of R. rhodochrous XplB.

Figure 5.3: Ni-NTA affinity and DEAE ion exchange chromatography purification of XplB.

Figure 5.4: Ni-NTA affinity chromatography purification of TF-XplB.

Figure 5.5: Ni-NTA affinity chromatography purification of thrombin treated TF-XplB.

Figure 5.6: Ni-NTA affinity and hydroxyapatite chromatography purification of XplB.

Figure 5.7: Multiple amino acid sequence alignment of the XplB with other flavodoxin/ferredoxin reductases.

Figure 5.8: CD spectra of XplB.

Figure 5.9: MALLS analysis of XplB.

Figure 5.10: NADPH-dependent XplB reduction.

Figure 5.11: FAD- and XplB-mediated XplA activity.

Figure 5.12: Quantification of the FMN content of XplA.

Figure 5.13: Reconstitution of XplA with FMN.

Figure 5.14: Analysis of FMN binding to XplA.

Figure 5.15: Analysis of FMN binding to XplA in NaCl-containing buffer.

Figure 5.16: Multiple amino acid sequence alignment of the XplB flavodoxin with other flavodoxins.

Figure 5.17: Determination of reduction potentials for the FMN cofactor in XplA.

Figure 5.18: Time-lapse production of XplA flavodoxin.

Figure 5.19: Ni-NTA affinity chromatography purification of XplA flavodoxin.

Figure A1: E. coli codon-optimised XplA DNA sequence

Figure A2: E. coli codon-optimised XplB DNA sequence

List of Tables

Table 2.1: Buffer compositions.
Table 2.2: JBS Solubility Kit part A: Buffer kit (100 mM).
Table 2.3: JBS Solubility Kit part B: Additive kit.
Table 4.1: X-ray data collection and refinement statistics for R. rhodochrous strain 11Y XplA-HD.
Table A1: UV/Visible absorption and binding affinity data for ligand binding to XplA.
Table A2: Melting temperatures of buffer screen of XplA-HD.
Table A3: Melting temperatures of additive screen of native and ligand-bound XplA-HD.

8
**Abbreviations**

δ-ALA  δ- aminolevulinic acid  
4-PIM  4-phenylimidazole  
AdR  Adrenodoxin reductase  
Adx  Adrenodoxin  
AMP  Adenosine monophosphate  
APS  Ammonium persulphate  
BB  BugBuster  
BME  β-mercaptoethanol  
BSA  Bovine serum albumin  
Bp  Base pair  
CD  Circular dichroism  
CO  Carbon monoxide  
CPR  NADPH-cytochrome P450 reductase  
CSS  Clear strategy screen  
CT  Charge transfer  
CYP  Cytochrome P450  
DEAE  Diethyl aminoethyl cellulose  
dH₂O  Distilled deionised water  
DMS  Dimethyl sulphide  
DMSO  Dimethyl sulfoxide  
DNA  Deoxyribonucleic acid  
dNTP  Deoxyribonucleotide triphosphate  
DTT  Dithiothreitol  
EDTA  Ethylenediaminetetraacetic acid  
EMS  Ethyl methyl sulphide  
EPR  Electron paramagnetic resonance  
FAD  Flavin adenine dinucleotide  
Fdx  Ferredoxin  
Fe-S  Iron-sulphur  
FldR  *E. coli* flavodoxin reductase  
FMN  Flavin mononucleotide  
FNR  Ferredoxin-NADP+ reductase  
FprA  Flavoprotein reductase A  
FUV  Far ultraviolet  
GR  Glutathione reductase  
GST  Glutathione S-transferase  
HCl  Hydrochloric acid  
HS  High spin  
IPTG  Isopropyl-β-D-thiogalactoside  
ITC  Isothermal titration calorimetry  
JCSG  Joint centre for structural genomics  
KBr  Potassium bromide  
Kₐ  Dissociation constant  
kDa  Kilo Dalton  
KPi  Potassium phosphate  
LS  Low spin  
LB  Luria Bertani  
MALLS  Multiangle laser light scattering  
MCD  Magnetic circular dichroism
MEDINA  Methylenedinitramine
$M_r$  Molecular mass
MWCO  Molecular weight cut off
NAD (NAD$^+$)  Nicotinamide adenine dinucleotide
NADH  NAD (reduced form)
NADP (NADP$^+$)  Nicotinamide adenine dinucleotide phosphate
NADPH  NADP (reduced form)
NDAB  4-nitro-2,4-diazabutanal
NHE  Normal hydrogen potential
NIR  Near infrared
NO  Nitric oxide
NMR  Nuclear magnetic resonance
NUV  Near ultraviolet
OD  Optical density
O/N  Overnight
PCR  Polymerase chain reaction
PDA  Photodiode array
PDB  Protein Data Bank
PdR  Putidaredoxin reductase
Pdx  Putidaredoxin
PEG  Polyethylene glycol
$pK_a$  Acid dissociation constant
PP  Pact Premier
RDX  Royal Demolition eXplosive (Royal Department X)
$R_h$  Hydrodynamic radius
R/T  Room temperature
SOC  Super optimal broth
SDS-PAGE  Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TB  Terrific broth
TAE  Tris/acetate/EDTA
TE  Tris/EDTA
TEMED  $N,N,N',N'$-tetramethylethylenediamine
TF  Trigger factor
TCB  Trichlorobenzene
TNB  Trinitrobenzene
TNT  Trinitrotoluene
UV  Ultraviolet
v/v  Volume to volume
w/v  Weight to volume
XplA-HD  XplA heme domain
XplA-FMN  XplA flavodoxin domain
YT  Yeast tryptone
Abstract

Hexahydro-1,3,5-trinitro-1,3,5-triazene (RDX) is a military explosive that has become a recalcitrant environmental pollutant over the last few decades owing to its production, storage and use. CYP177A1 (XplA) is a biotechnologically interesting and novel class of P450-flavodoxin fusion enzyme identified from *Rhodococcus rhodochrous* strain 11Y that catalyses the breakdown of RDX. Its redox partner is a NAD(P)H-dependent FAD-binding flavodoxin reductase (XplB). This study reports the biochemical, biophysical and structural properties of these two enzymes which form a novel P450 redox system with unique domain organisation. These reveal novel features for a P450 enzyme with non-standard UV/Visible spectroscopic features and unusual ligand binding properties. Unexpectedly, XplA’s affinity for imidazole is exceptionally high ($K_d = 1.57 \mu M$), explaining previous reports of a redshifted XplA Soret band in pure enzyme. XplA’s true Soret maximum is at 417 nm. Similarly, the XplA flavodoxin domain displays unusually weak FMN binding ($K_d = 1.09 \mu M$), necessitating its reconstitution with the FMN cofactor. Ligand binding data demonstrate XplA’s constricted active site, which can only accommodate RDX and small inhibitory ligands (e.g. 4-phenylimidazole and morpholine) while discriminating against largerazole drugs. The crystal structure identifies a high affinity imidazole binding site, consistent with its low $K_d$, and shows active site penetration by PEG, perhaps indicative of an evolutionary lipid metabolising function for XplA. The substrate-free heme iron potential (-268 mV vs. NHE) is positive for a low spin P450, consistent with the predominantly reductive role of XplA. The elevated potential of the FMN semiquinone/hydroquinone couple (-172 mV) is also consistent with this functional adaptation. The XplB reductase partner could not be isolated with the FAD cofactor incorporated to make holoprotein. However, the protein was isolated in a soluble and homogenous state which demonstrated very weak FAD affinity. XplB’s ability to interact with XplA and a pyridine nucleotide coenzyme was demonstrated, indicating the enzyme was functional in the presence of FAD. XplA’s unusual molecular selectivity, structural and thermodynamic properties likely reflect its evolution as a specialised RDX reductase catalyst.
Declaration

The author declares that no part of the work presented in this thesis has been submitted in support of an application for another degree or qualification in 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 for any 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, trademarks 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 firstly like to thank Professor Andrew Munro for his guidance and encouragement, and secondly for giving me the opportunity to do a PhD, something I had never previously given consideration and thought capable, and also for the opportunity to work and be part of his wonderful laboratory and group. This group is full of fantastic people who have all given me much needed advice and support throughout my PhD, and contributed greatly to my development as a scientist and also as a person. Many have come and gone in my time, but their kindness and generosity shall not be forgotten, and all shall be fondly remembered as part of the Munlops, to which I proudly belonged.

I must give special thanks to Dr Kirsty McLean and other senior members of the group, who had the unfortunate task of weaning me in and showing me the ropes, and also to Marina Golovanova for all her help over the last three years and for putting up with my cheekiness and disobedience.

I would also like to extend my thanks to members of the extended Molecular Enzymology group, all of whom have at some point been a source of help, and also friendship. The wonderful chemistry in the lab, to which each individual contributed, has made my time and importantly my work so much more enjoyable.

Finally I would like to thank my family for supporting my decision to do a PhD, and providing encouragement to pursue a career in science.
Chapter 1: Introduction

1.1: Cytochromes P450

Cytochromes P450 are a superfamily of heme $b$-containing monooxygenase enzymes containing protoporphyrin IX as their prosthetic heme factor and first described by Mason in 1955 and Hayaishi in 1957 [1, 2]. Porphyrins are large heterocyclic organic rings formed by four interconnected modified pyrrole molecules. Binding of a central iron produces the heme cofactor, and further modifications of the pyrrole molecules give rise to different types of hemes that can be incorporated into a variety of proteins, the names of which are usually determined by the heme they carry, e.g. cytochrome $a$, and cytochrome $c$. The structures of three common hemes incorporated by proteins are shown in Figure 1.1.

![Figure 1.1: Chemical structures of heme a, b and c. Common heme prosthetic groups incorporated by proteins. The central iron is bound to four nitrogens of the pyrrole rings. Different substituents at positions 3, 8 and 18 give rise to the various hemes, and other minor heme groups are generally further derivatives of the major hemes shown. Heme $b$, which is utilised by P450s is the most common heme found in proteins. In P450s the iron is also coordinate bonded to a conserved cysteine residue and to a water molecule to form a hexa-coordinated state. Structures were drawn using ChemDraw [3].](image)

Interest in P450s began due to the importance of these catalysts in human physiology and pharmacology, and also due to their unique ability to perform oxidative chemistry [4-6]. Most of our current understanding of the biochemical, biophysical and structural properties of P450s derive from two well characterised P450s, CYP101A1 (P450cam) from *Pseudomonas putida* and CYP102A1 (BM3) from *Bacillus megaterium* [7-9]. There are now more than 14,000 P450 genes identified,
representing one of the largest superfamilies of enzymes [10]. P450s are grouped into families depending on the reactions they catalyse, with P450s sharing 40% amino acid sequence identity grouped into the same family [11]. Furthermore, those with a sequence identity over 55% are grouped into sub-families. The classification and nomenclature of P450 enzymes is maintained by David Nelson at the University of Tennessee [10].

In the resting state of the enzyme the heme iron is in the ferric (III) oxidation state, and is normally ligated to a pair of axial ligands forming hexa-coordinated heme (Figure 1.2A). One of these is an absolutely conserved cysteine residue which is the proximal ligand. This cysteine residue, which can be in the thiol (SH) or thiolate (S-) state, and its interactions with neighbouring residues, is responsible for the unique chemical properties of the P450s [12-15]. The other axial ligand is typically water [16]. The name cytochromes P450 is derived from the fact that they are coloured proteins, having a “pigment at 450 nm”, with a characteristic absorption maximum called the “Soret” peak formed by absorbance of light at wavelengths near 450 nm when the heme iron is reduced to the ferrous (II) state and complexed to carbon monoxide (CO), as was first described by Omura and Sato in 1964 (Figure 1.2B) [17, 18].

![Coordination arrangement of hema-coordinated P450 heme in the resting and CO-bound states.](image)

**Figure 1.2:** Coordination arrangement of hexa-coordinated P450 heme in the resting and CO-bound states. In the resting state, the heme iron is in the ferric state with water and cysteine thiolate as the axial ligands. Reduction of the heme produces ferrous heme, which can bind CO to form a Fe(II)CO adduct, which can exist either as the P450 complex when the cysteine is in the thiolate state, or the P420 complex if the cysteine is protonated [19, 20]. The P420 complex is characterised by a Soret peak near 420 nm, as opposed to the P450 complex which typically has a Soret peak near 450 nm. Structures were drawn using ChemDraw, and adapted from [21].
Chapter 1: Introduction

1.1.1: Physiological Roles of P450s

P450s are ubiquitous in nature with a diverse range of physiological roles. They are found in all domains of life and are involved in a host of metabolic pathways in eukaryotes and prokaryotes [22-24]. Subsequently, P450s catalyse a diverse range of chemical reactions including hydroxylation, epoxidation, demethylation, dehalogenation, desaturation and isomerisation [25, 26]. The majority of P450s catalyse the oxidation of a large range of organic molecules by reductive breakage of the dioxygen bond, leading to the introduction of a single atom of oxygen into an organic substrate and the production of a single molecule of water. The reaction is summarised below in Equation 1.1 (and elaborated further in section 1.1.1.3), where RH is the substrate, and ROH is the hydroxylated product. The electrons are supplied by a pyridine nucleotide coenzyme.

Equation 1.1:

\[ RH + O_2 + NAD(P)H + H^+ \rightarrow ROH + H_2O + NAD(P)^+ \]

Although all P450s share the ability to perform oxidative chemistry, they can be separated by their physiological role in their respective organisms, e.g. P450s can function as hydroxylases, demethylases and epoxidases [27-29]. Some of these organisms utilise P450s that have evolved to perform other functions that do not involve oxygen activation/insertion, e.g. P450s that have evolved to function as reductases [30].

Bacterial and mammalian microsomal P450s have been extensively studied, but in recent years interest has also grown for P450s in insects and plants. Plant P450s likely constitute the largest complement of P450s, with *Arabidopsis thaliana* (thale cress) alone containing 249 genes coding for P450s [31], and *Oryza sativa* (rice) containing 323 P450 (CYP) genes, respectively [32]. The large repertoire of P450s carried by plants likely reflects their sessile nature, which necessitates the use of novel methods to avoid predation and acquire nutrition. Plant P450s are known to be involved in the synthesis of many chemicals that are toxic to pests and herbivores, e.g. *Apium graveolens* (celery) produces furanocoumarins to ward off herbivores [33]. The herbivore in turn, *Helicoverpa zea* (corn earworm) is able to counteract the toxin in a P450-dependent manner [34]. Many insect P450s are also known to have similar
roles in detoxifying insecticides [35]. P450s are also involved in other important physiological roles in these organisms, including the production of juvenile insect hormones [36], and in synthesis of a vast number of organic products in plants [37]. However, the roles of P450s in many of these processes remain unclear.

1.1.1.1: Mammalian P450s

Mammalian P450s are best known as detoxification enzymes found in the endoplasmic reticulum membrane of liver microsomes that catalyse the breakdown of drugs and other exogenous compounds [38, 39]. Typically, they catalyse the oxidation of lipophilic compounds to produce more hydrophilic compounds to expedite their removal from the body. Microsomal P450s have been described as “generalist” enzymes as the substrates they recognise are generally more diverse than their bacterial counterparts, and many are of pharmacological relevance [40, 41], e.g. CYP2B6 has been shown to be involved in the metabolism of bupropion and ifosfamide, which is used in treatment for depression and attention-deficit/hyperactivity disorder, and for treatment of cervical cancer, respectively [42, 43].

In addition to the microsomal P450s, there are also P450s located in mitochondria [44-46]. These are primarily involved in steroid biosynthesis, e.g. CYP11A1 (P450scC) is able to catalyse three consecutive reactions involving two hydroxylations followed by a C-C bond scission in cholesterol side chain cleavage to produce pregnenolone, a precursor of many important steroid hormones [47].

Mammalian P450s are notoriously difficult to characterise due to their membranous nature, which makes them difficult to express and purify. The first mammalian P450 to be successfully expressed was CYP2A1 in 1989 by Asseffa et al. using baculovirus [48]. More recently, mammalian P450s have been expressed as truncated constructs missing their N-terminal anchor regions [49-51]. The absence of this hydrophobic anchoring sequence greatly increases the solubility of these P450s, making them more amenable to heterologous recombinant protein expression using Escherichia coli expression systems. The best characterised of these is CYP3A4, which is responsible for the metabolism of ~50% of marketed drugs [52, 53]. Progress culminated in the determination of the crystal structure of CYP2C5 by Williams et al. in 2000, the first crystal structure of a mammalian P450 [54]. This has paved the way
Chapter 1: Introduction

for the characterisation of numerous mammalian P450s, which are attractive to pharmaceutical companies because of their role in drug metabolism, and also due to their generalist nature, which makes them more amenable to manipulation to catalyse the oxidation of other substrates to produce high value compounds [55, 56].

1.1.1.2: Bacterial P450s

Although the roles of P450s in eukaryotic organisms have long been defined, their roles in prokaryotes are less well defined. Bacterial P450s are normally involved in the catabolism of exogenous organic compounds to utilise as a source of nutrition. Their substrate specificity can be described as “specialist”, and as opposed to microsomal P450s, their active sites are not as flexible [40]. However, it is becoming increasingly clear that bacterial P450s also have important physiological roles, e.g. CYP107H1 (P450Biol) is involved in biotin synthesis in Bacillus subtilis [57]. In addition, many bacterial P450s are of environmental and medical relevance, e.g. the actinobacterial species produces a wealth of natural products, many of which are antibiotics [58-62].

Importantly, bacterial P450s are much more soluble than microsomal P450s and as such, their properties are better understood and the characterisation of numerous bacterial P450s have laid the foundations for understanding P450 biochemistry. The soluble nature of bacterial P450s is mainly due to the lack of an N-terminal peptide region that anchors eukaryotic P450s to cellular membranes. The first bacterial P450 to be characterised was P450cam, which catalyses the hydroxylation of camphor to enable its metabolism for energy in P. putida [8, 63]. Subsequent work by Narhi et al. on BM3, a fatty acid hydroxylase, has played a crucial role in P450 research because it has been used as a model for eukaryotic P450s, due to its similarities to eukaryotic P450s and since it was the first bacterial P450 found to have a eukaryotic-like NADPH-cytochrome P450 reductase (CPR) redox partner enzyme fused to the P450 [9].

1.1.1.3: Catalytic Cycle

Understanding of the mechanism of the P450 catalytic cycle has come from the direct observation of intermediates through a variety of spectroscopic techniques, and the use of appropriate substrates with mechanistically revealing rearrangements during oxidation. The elucidation of the catalytic cycle of P450s was mainly carried out on
P450cam and more recently on CYP119 from a thermophilic bacterium, and is schematically depicted in Figure 1.3 [64-66].

Figure 1.3: Schematic representation of the cytochrome P450 catalytic cycle. Binding of the oxidisable substrate (RH) starts the catalytic cycle by displacing the water ligand. This facilitates delivery of the first electron, followed by binding of dioxygen to form a ferrous dioxygen complex (4), which accepts the second electron to form a ferric-peroxo species (5). Two further protonation steps produces the iron-oxo intermediate (Compound I) (7), and concomitant production of a molecule of water. Insertion of the oxygen atom into the substrate, subsequent release of the product and rebinding of a water molecule to the ferric heme iron regenerates the resting enzyme. The reaction indicated by the diagonal arrow is the "peroxide shunt" in which addition of hydrogen peroxide (or an organic peroxide) can be used to surrogate for electron delivery from a redox partner (along with proton and oxygen delivery), producing ferric hydroperoxide (Compound 0) (6) directly [67]. However, this is generally an inefficient means of performing P450 catalysis. Scheme was drawn using ChemDraw and Microsoft PowerPoint.

The resting state of the enzyme is a ferric (III) complex (1) in the low spin (LS) state that is hexa-coordinated with water and thiolate axial ligands as described earlier.
Chapter 1: Introduction

Binding of the substrate displaces the water molecule giving rise to a pentacoordinated heme species, which promotes the iron spin state change from LS to high spin (HS) (2). This shift is accompanied by an increase in the redox potential of the heme iron (Fe\(^{3+}\)/Fe\(^{2+}\)) and it receives an electron from a redox partner that reduces (2) to the HS ferrous (II) complex (3). Ferrous iron porphyrin is a good dioxygen binder, and leads to the binding of molecular oxygen to produce the LS ferrous-dioxygen complex (4) (oxyferrous). This complex is another good electron acceptor, possibly due to the electron withdrawing properties of the oxygen (from the iron), and receives another electron to give rise to the twice-reduced ferric-peroxo species (5). The ferric-peroxo complex undergoes protonation to yield the ferric hydroperoxide complex (6) commonly referred to as Compound 0. This compound subsequently undergoes a second protonation and releases a water molecule leading to the high-valent iron-oxo complex also known as Compound I, which was first characterised in P450s by Rittle et al. [66]. This complex finally transfers the distal oxygen atom to the substrate, which is released and is replaced by a water molecule to regenerate the resting state of the enzyme.

Compound I, which does not accumulate under normal turnover conditions, is generally accepted to be the active species which attacks the substrate [68-72]. The later reactive intermediates in the catalytic cycle are transient and until recently compelling structural and spectroscopic evidence for their formation and reactivity was difficult to attain [73, 74]. However, some of these later reactive intermediates are more stable in other thiolate-ligated hemoproteins like chloroperoxidases (CPO), which have intermediates similar to their P450 counterparts. CPO Compound 0 and Compound I have been characterised and have served as a valuable model system for P450s [75-79]. P450 and CPO Compound I share similar Mössbauer and UV/Visible spectroscopic signatures, but they have different electron paramagnetic resonance (EPR) signatures [66]. Some of the later reactive intermediates preceding Compound I are also capable of attacking the substrate [80-82], e.g. Compound 0 of CYP161A2 (P450 PimD) from *Streptomyces natalensis* catalyses the epoxidation of de-epoxypimaricin to pimaricin [83, 84].

The mechanism of oxygen insertion into the substrate is generally accepted to proceed through the hydrogen abstraction oxygen rebound mechanism proposed by Groves in 1978 (Figure 1.4) [85]. The high-valent Compound I attacks the R-H bond...
Chapter 1: Introduction

and abstracts the hydrogen from the substrate to form a ferryl-hydroxyl species (Compound II), which essentially remains uncharacterised, and a substrate radical. The substrate radical then attacks Compound II and forms a bond with the hydroxyl group to produce a hydroxylated product and restoring the heme iron to the ferric state [86-88].

Figure 1.4: Proposed hydrogen abstraction and oxygen rebound mechanism of the P450s. Compound I (left) extracts a hydrogen from the substrate by heterolytic cleavage of the R-H bond and forms a protonated ferryl-oxo intermediate (Compound II) and a substrate radical. The substrate radical attacks Compound II, forming a bond with the hydroxyl group to produce the hydroxylated product and regenerating ferric heme. Scheme was drawn using ChemDraw, and adapted from [89].

The P450 catalytic cycle is self-regulatory, and in most P450s the oxidised Fe(III) has a relatively low redox potential, compared to other hemoproteins, e.g. myoglobin has a redox potential of +46 mV [90]. This is due to the cysteinate ligand, which exerts an electronic ‘push’ effect on the heme iron [12, 91]. This has the effect of increasing the electron density of the heme iron, reducing its propensity to accept electrons. Furthermore, studies have shown that the redox potential of the heme iron is influenced by its exposure to the solvent [92, 93]. Substrate binding decreases the solvent exposure of the heme by displacing water molecules from the active site. Substrate binding also displaces the 6th axial water ligand, leading to the spin state change. The change in spin state has the effect of increasing the cationic radius of the Fe(III), and to accommodate this increased size it moves out of the plane of the porphyrin moiety [94, 95]. These changes increase the redox potential of the heme iron. This mechanism prevents wasting reducing equivalents in the absence of substrate.
Chapter 1: Introduction

The concomitant spin state shift that accompanies substrate binding can be explained by ligand field theory, a characteristic of coordination complexes. In the resting state where the heme is in an octahedral geometry with six ligands, the ligands introduce an energy difference between the $d$-orbitals of the heme iron, the octahedral splitting energy ($\Delta_{\text{oct}}$), leading to two groups of $d$-orbitals ($e_g$ and $t_{2g}$) (Figure 1.5) \[96, 97\]. The degree of splitting is dependent on the ligand, and the field strength generated by the ligand is determined by the spectrochemical series \[97\]. Strong field ligands, which are $\pi$-acceptors such as CO and CN$^-$, increase the energy difference between low energy orbitals and high energy orbitals, so that it becomes energetically more expensive to place an electron in a high energy orbital than it is to spin pair. Consequently, binding of these ligands typically produces LS heme. Conversely, low-field ligands are all $\pi$-donors like I$^-$, and produce a small $\Delta_{\text{oct}}$, which favours the HS state. Furthermore, in the absence of the aqua 6$^\text{th}$ ligand, which is a medium field ligand, the $\Delta_{\text{oct}}$ is less than the spin pairing energy and some of the electrons will be in the high energy orbitals leading to HS heme.

Another important aspect of the P450 catalytic cycle is the mechanism and pathway of proton delivery to the heme iron, which remains an active area of research. It is thought to involve networks of water molecules and amino acid side chain hydroxyl groups \[98, 99\]. These water molecules and amino acids play an important role in the P450 catalytic cycle by maintaining rapid and timely delivery of electrons and...
Inefficient electron or proton delivery to the heme iron can lead to uncoupling of oxygen activation and redox partner oxidation with the subsequent release of hydrogen peroxide, superoxide or water according to the point in the catalytic cycle reached \([21, 100]\). One particular residue that is almost absolutely conserved in P450s whose primary function is to catalyse the more traditional P450 reactions, such as hydroxylations, is a threonine residue demonstrated to be important for efficient catalysis \([101, 102]\). The corresponding residue in P450cam is T252, which together with D251 are referred to as the acid/alcohol pair that plays a critical role in the hydrogen bonding network essential for delivery of protons to the active site \([101]\). T252 is in hydrogen bonding distance to the bound dioxygen and to a water molecule \([70]\). Furthermore, T252 promotes the addition of the second proton to the distal oxygen by accepting a hydrogen bond from the hydroperoxy intermediate \([99]\). In P450 oxygenases that do not contain the threonine residue, it has been shown that a substrate hydroxyl group takes part in the hydrogen bonding network in place of the missing threonine \([98, 103]\). Mutagenesis of this residue produces mutants with low coupling efficiency between NAD(P)H oxidation and substrate hydroxylation \([104, 105]\).

Although it was proposed earlier that all P450s are capable of catalysing oxygen activation chemistry, there are P450s that do not fully follow this catalytic cycle and some P450s do not appear to be involved in oxygen activation/insertion, e.g. CYP177A1 (XplA) from *Rhodococcus rhodochrous* \([106]\). Furthermore, P450s can catalyse monooxygenation reactions in the absence of molecular oxygen when provided with organic or inorganic peroxides using the peroxide shunt mechanism and directly forming Compound 0 \([67, 107]\). Moreover, the peroxide shunt mechanism does not require NAD(P)H, O\(_2\) or any redox partners, e.g. CYP152A1 (P450 Bsb) from *B. subtilis* is a naturally H\(_2\)O\(_2\) driven P450, catalysing \(\alpha\) and \(\beta\) hydroxylation of fatty acids using H\(_2\)O\(_2\) as the oxidant \([107]\).

### 1.1.2: P450 Structure

As briefly discussed previously, the first P450 to be characterised in detail was P450cam, and this was also the first P450 to have its crystal structure determined \([108]\). Much of the current understanding of the structure-function relationships in
Chapter 1: Introduction

P450s is based on studies of P450cam [109, 110]. Since that breakthrough, numerous P450 structures have been determined, e.g. the heme domain of BM3 (BMP) [111]. The P450 fold can be described as “prism” shaped, and it appears that the overall P450 fold is quite well conserved throughout the P450 superfamily, and unique to the P450s (Figure 1.6). The heme is sandwiched between two distinct domains, a small β-sheet rich domain and a larger α-helix rich domain.

Although the overall fold is maintained, the precise positioning of various structural elements differs substantially. In general, the closer to the heme the more conserved the structure, especially for helices I and L which directly contact the heme [105]. Conversely, the regions that control substrate specificity display the greatest variation, especially in the B’ helix. The most conserved elements of the P450 structure are those involved with the heme-thiolate oxygen activation chemistry, including the cysteine residue discussed earlier. This cysteine residue is part of the heme binding motif highly conserved in P450s. The motif is located either before or at the start of the L helix and contains the consensus sequence FxxGxxxCxG, where x is any residue. The motif also contains a phenylalanine residue, which has an important role in modulating the thermodynamic properties of the heme iron [14, 112]. In addition to the highly conserved acid/alcohol pair discussed earlier, a glycine (G248 in P450cam) is also well conserved and serves as a hydrogen bond donor to T252, and is almost exclusively conserved as a small residue in proximity to heme-bound dioxygen [70, 99].
Figure 1.6: **P450 structures illustrating the common three-dimensional fold.** The P450 fold is composed of a minor β-sheet rich domain (magenta), and a major α-helix rich domain (orange), and can be tentatively described as prism shaped. The latter domain contains the heme prosthetic group, shown in red as a stick model, and the I helix in blue, which spans the entire structure and contains many of the conserved residues required for efficient catalysis. P450s shown are P450cam (PDB code; 2CPP) from *P. putida* with all major helical segments labelled, BMP of BM3 (PDB code; 2IJ2) from *B. megaterium* and CYP119 (PDB code; 1I07) from *Sulfolobus solfataricus*. Structures were drawn using PyMol [113].
Since their discovery researchers have been most intrigued by how P450s were able to catalyse the monooxygenation reactions, and the deciphering of the catalytic cycle was of great help in this regard. Their ability to oxidise unactivated C-H bonds has also captivated the interest of chemists for five decades. Furthermore, P450s showed tremendous industrial potential due to their great versatility in catalysing a diverse range of reactions [58, 114, 115]. In particular, P450s have tremendous potential in the production of high value compounds, due to their ability to catalyse regio- and stereo-specific hydroxylations in contrast to traditional methods of organic synthesis [55, 56, 116]. However, despite their potential and the proposal of numerous biotechnological applications of P450s, their use for biotechnological purposes has thus far been limited by numerous factors, including instability of the enzymes and low catalytic activity [117-119]. P450s also require a continuous supply of electrons, which are normally derived from NAD(P)H, a costly cofactor, which has hindered the use of P450s in industrial processes. The peroxide shunt mechanism has been pursued as an alternative source of activated oxygen, and the discovery of novel P450-redox partner fusion enzymes has removed the need for redox partners [120, 121]. However, the use of the peroxide shunt mechanism has been hindered by rapid inactivation of enzymes by peroxide and other reactive oxygen species.

Due to the reasons described above, whole cell approaches have been favoured compared to isolated P450 enzyme systems, e.g. in the synthesis of drugs such as pravastatin from compactin by P450sca-2 in E. coli and Streptomyces carbophilus [116, 122], and pregnenolone from ergosterol in Saccharomyces cerevisiae expressing bovine P450scc [123]. Other successful applications of P450s in industrial processes have been reported, famously in the production of blue roses by introducing the cyp75a gene, a flavonoid hydroxylase, into transgenic plants [124, 125]. Fungal CPOs and bacterial P450s have also been genetically engineered to do large scale biotransformations [126-128].

More importantly, P450s have also been successfully utilised in the detoxification of environmental pollutants, notably in phytoremediation, e.g. several human P450s (CYP1A1, CYP2B6, CYP2C9 and CYP2C19) have been introduced into various plants for the effective removal of herbicides, [129, 130]. The use of the flavoprotein pentaerythritol tetranitrate reductase (PETNR) in the bioremediation of 2,4,6-
trinitrotoluene (TNT) using transgenic tobacco plants has also demonstrated the applicability of using recombinant enzymes for bioremediation [131, 132]. P450s also show potential in removal of three major classes of environmental pollutants, polycyclic aromatic hydrocarbons (PAHs), polychlorinated dibenzo-p-dioxins (PCDDs) and polychlorinated biphenyls (PCBs) [133-136].

Great effort has been put into engineering P450s, with BM3 in particular receiving copious attention owing to its fast rate of catalysis [137-139]. Several approaches have been employed for the genetic engineering of P450s, either to improve existing properties, or to introduce novel functional capacity by altering substrate specificity, depending on the level of information available about the natural enzymes. This is achieved either by rational design using site-directed mutagenesis of an enzyme with a well defined structure-function relationship, or by directed evolution through random mutagenesis when only limited information is available on the mechanism of the enzymes to identify mutants with a desired activity or property [140, 141].

1.2: Flavoproteins

The flavoproteins consist of a diverse range of proteins containing a riboflavin (vitamin B2) derivative as their cofactor. The presence of the bound flavin extends the catalytic ability of these proteins to catalyse redox reactions, creating versatile proteins that have diverse cellular functions, including oxygen activation, light sensing and protein folding [142-144].

Riboflavin consists of an isoalloxazine ring bound to a ribitol sugar at the N10 position. Phosphorylation of the terminal hydroxyl group of the ribityl tail by riboflavin kinase produces flavin mononucleotide (FMN) [145]. Adenylation of FMN by FMN adenyltransferase produces flavin adenine dinucleotide (FAD) [146]. The flavin cofactor found in flavoproteins is normally FMN or FAD, although flavins with further modifications are also incorporated by flavoproteins [147-149]. The oxidised isoalloxazine ring imparts a yellow colour on the molecule. The fully reduced form is colourless. Flavins also show strong fluorescence due to the isoalloxazine ring [150]. The structures of all three molecules are shown in Figure 1.7.
Chapter 1: Introduction

Figure 1.7: Chemical structures of riboflavin and its derivatives FMN and FAD.
Riboflavin consists of the flavin moiety formed by the isalloxazine ring, linked at the N10 position to the reduced form of a ribose sugar. Phosphorylation of the ribityl tail at the terminal hydroxyl group position produces FMN, and adenylation of this compound produces FAD. In these structures, only the *si*-face of the isalloxazine ring is shown, the reverse side of the ring is known as the *re*-face. Structures were drawn using ChemDraw.

The versatility of the flavoproteins is due to the unique ability of the flavin to exist in three redox states, oxidised (OX), one-electron reduced (semiquinone (SQ)) and two-electron reduced (hydroquinone (HQ)), allowing them to catalyse single or two electron reductions (Figure 1.8) [151]. The small iron-sulphur (Fe-S) cluster ferredoxins for example are only capable of catalysing single electron reductions, and pyridine nucleotides are only capable of catalysing two-electron transfers via the hydride ion [152]. In addition, the SQ and HQ species can exist in unprotonated and protonated states leading to the anionic or neutral species. The anionic SQ displays a red colour, and the neutral SQ is blue. The SQ of the free flavin disproportionates to the oxidised and fully-reduced forms [153].

The versatility of the flavoproteins is due to the unique ability of the flavin to exist in three redox states, oxidised (OX), one-electron reduced (semiquinone (SQ)) and two-electron reduced (hydroquinone (HQ)), allowing them to catalyse single or two electron reductions (Figure 1.8) [151]. The small iron-sulphur (Fe-S) cluster ferredoxins for example are only capable of catalysing single electron reductions, and pyridine nucleotides are only capable of catalysing two-electron transfers via the hydride ion [152]. In addition, the SQ and HQ species can exist in unprotonated and protonated states leading to the anionic or neutral species. The anionic SQ displays a red colour, and the neutral SQ is blue. The SQ of the free flavin disproportionates to the oxidised and fully-reduced forms [153].
Figure 1.8: **Chemical structures of the three redox states of FMN.** (F) Oxidised flavin, when reduced ($E_1 = -238$ mV) can form the neutral blue (FH) or anionic red SQ (F\textsuperscript{−}). A second reduction step ($E_2 = -172$ mV) fully reduces the flavin to either a neutral (FH\textsubscript{2}) or anionic (FH\textsuperscript{−}) HQ. Redox potentials and the pK\textsubscript{a} values shown are for free FMN. Both properties are highly influenced by the protein environment [154, 155]. Structures were drawn using ChemDraw, and adapted from [151].

The protein environment modulates the thermodynamic properties of the bound flavin by differentially stabilising the three redox states [154, 155]. Furthermore, the binding of pyridine nucleotides normally increases the redox potential of the flavin such that electron transfer is energetically favourable [156, 157]. The mechanisms are not very well understood, but likely involve subtle structural changes in the flavin environment induced by coenzyme binding that increases the stability of the SQ or HQ with respect to the oxidised flavin.

Although the majority of flavoproteins contain non-covalently bound flavin cofactors, some flavoproteins have evolved to incorporate a covalently attached flavin, e.g. two flavoproteins from *Vibrio cholera*, Rnf and RnfD, have both been shown to covalently bind FMN and flavin moiety, respectively, through threonine residues [158]. The nature of the flavin covalently bound to RnfD was not confirmed. FAD is more
commonly covalently attached to proteins, e.g. in succinate dehydrogenase, where the FAD is covalent bound to a histidine through the 8-methyl group of the isoalloxazine ring [159]. The role of covalent linkage of flavins to proteins is not well understood, and may well influence the redox potential of the flavin, or enhance structural integrity [160, 161].

Classification of the flavoproteins has historically been difficult due to structural diversity, and there are at least four generally accepted classes of FAD-binding flavoproteins, sharing common topological features such as the Rossmann fold [162]. However, only the ferredoxin-NADP+ reductase family (FNR) and the FMN-binding flavodoxin proteins, both being integral components of the P450 redox system, will be introduced here. These redox proteins have been extensively characterised owing to their ease of expression, purification, stability and spectroscopic properties imparted by the yellow flavin cofactor.

1.2.1: Ferredoxin-NADP+ Reductases

FNRs interact with pyridine nucleotides and ferredoxins in electron transfer pathways, famously in photosynthesis where FNR oxidises two molecules of reduced-ferredoxin and reduces NADP+ to regenerate NADPH [163]. They are also involved other important processes such as nitrogen fixation and oxidative stress responses [164-167]. FNRs are able to catalyse one and two electron exchange reactions, and in the P450 redox system they are more noted for catalysing the reverse reaction, where they receive two electrons from a pyridine nucleotide coenzyme via a hydride ion transfer and release the electrons one at a time to an electron acceptor. FNR was first isolated by Avron et al. from pea thylakoids and, as mentioned above, are one of four FAD-binding protein families identified by Dym and Eisenberg [168, 169]. Interactions between the FNR and NADPH are thought to involve aromatic stacking of the nicotinamide and isoalloxazine rings, and electrostatic interactions between the 2’-phosphate group of the adenosine monophosphate (AMP) moiety with positively charged residues on the protein surface, including a conserved arginine [170, 171]. In addition, nuclear magnetic resonance (NMR) spectroscopic studies indicate that the pro-R-hydrogen attached to the C4 atom of NADPH is transferred directly to FAD as a hydride ion [172].
Enzymes in the FNR family vary in shape and size, with enzymes ranging from 29 kDa (*E. coli* flavodoxin reductase (FldR)) to 51 kDa (adrenodoxin reductase (AdR) from bovine mitochondria) [173, 174]. Although AdR displays FNR activity, it is more structurally related to the glutathione reductase family of FAD-binding proteins [175]. Other members of the glutathione reductase family that also display FNR-like activity include flavoprotein reductase A (FprA) from *Mycobacterium tuberculosis* [176]. Most FNRs share a common structure with spinach ferredoxin reductase made of two domains, an N-terminal FAD-binding domain and a C-terminal NADP+ binding domain (Figure 1.9). The latter domain comprises a variant of the Rossmann fold typically with a five-stranded central parallel β-sheet interconnected by α-helices. The Rossmann fold was first identified in 1974 by Michael Rossmann as a protein structural motif that binds nucleotides, notably NAD(P)H [162]. The fold is composed of six parallel β-sheets separated by α-helices, which form two symmetrical βαβαβ motifs interconnected by the third α-helix. Each motif constitutes a Rossmann fold that can bind one nucleotide, thus pairs of Rossmann folds are necessary to bind dinucleotides. The FAD-binding domain is a β-domain with a flattened six-stranded antiparallel β-barrel organised into two orthogonal sheets (β₁β₂β₅ and β₄β₃β₆) separated by one α-helix [177, 178].

**Figure 1.9:** *Structures of ferredoxin reductases.* The FAD-binding domain (blue) of AdR, which belongs to the glutathione reductase family, utilises a Rossmann fold distinct from FNRs, and the NADP+ binding domain (red) adopts a classical Rossmann fold [162]. FNRs also contain a FAD-binding domain (blue), and a NADP+ binding domain (red). The latter contains a modified Rossmann fold with a central five stranded parallel β-sheet. The FAD cofactors and NADP+ coenzymes are shown as stick models. Structures shown are bovine mitochondrial AdR (PDB code; 1CJC) and FNR from *Anabaena* (sp. pcc 7119) (PDB code; 1GJR) [171, 174]. Structures were drawn using PyMol.
Chapter 1: Introduction

1.2.1.2: FAD Binding

The FAD is bound through non-covalent interactions with the isoalloxazine ring, ribityl tail and pyrophosphate moieties. Many of the residues that play important roles in FAD and NAD(P)H binding are located in three conserved sequence motifs [179-181]. One of the most conserved sequence motifs is RxYS(T), where residues in brackets denote an alternative residue, which contains one of the two aromatic residues critical for FAD binding. The phenol ring of the conserved tyrosine stacks on the si-face of the isoalloxazine ring and also hydrogen bonds to the ribityl 4'-hydroxyl shown in Figure 1.10. Furthermore, the conserved arginine residue provides a hydrogen bond to the phosphate group [181]. Substitution of this arginine in rat CPR (R454) leads to a decrease in affinity for FAD by 25,000-fold [182]. The re-face of the isoalloxazine ring normally stacks with another aromatic residue, e.g. W677 in rat CPR is stacked against the re-face [183]. Substitution of either aromatic residue with a non-aromatic residue significantly decreases the affinity for FAD [182]. The aromatic residue stacked against the re-face of the isoalloxazine ring is not always present, and in FNRs from Pseudomonas aeruginosa and Azotobacter vinelandii, it is replaced by an alanine [178, 183]. The absence of this aromatic residue is likely to be functionally significant and favour NAD(P)H binding and accelerate the rate of FAD reduction.

The second conserved motif forms the pyrophosphate binding motif with the consensus sequence GxxS(T) [178]. The conserved alcohol residue forms a hydrogen bond with the phosphate group [183]. The position of the conserved glycine is thought to be key to binding FAD or FMN, and in FAD-binding proteins, this glycine is usually two or three residues preceding the alcohol residue [179]. The residues forming bonds with the pyrophosphate are the most conserved compared to residues involved in binding the adenine, ribitol and isoalloxazine moieties [169].

The third conserved motif with the consensus sequence GxGxxP is located in the NADP*-binding domain. Residues in this motif interact with the pyrophosphate group of the NAD(P)H [181, 184]. In addition, FNRs that utilise NADP* contain a S(T)R motif. The serine/threonine and arginine residues interact with the 2'-phosphate of the adenosine ribose moiety of NADP*. NAD*-binding oxidoreductases have an aspartate in place of this conserved alcohol residue [185].
Chapter 1: Introduction

Figure 1.10: FAD binding in E. coli FldR. The isoalloxazine ring of FAD (orange) is sandwiched between two aromatic residues (Y52 and Y247 (cyan)), and R50 forms a hydrogen bond with the pyrophosphate group. The FAD molecule adopts a bent conformation with the AMP moiety folded back and in proximity to the isoalloxazine ring, and forms favourable π-π interactions with W248. FAD binding can also adopt an elongated conformation, when the corresponding aromatic residue to W248 is absent. To enable electron transfer from NADPH, both Y247 and W248 must move away to expose the isoalloxazine ring. Structure (PDB code; 1FDR) was drawn using PyMol, and adapted from [181].

1.2.2: Flavodoxins

Flavodoxins are small electron transfer proteins ubiquitous in prokaryotes containing one molecule of non-covalently, but tightly bound FMN as its redox centre [186]. They are involved in a variety of microbial metabolic pathways catalysing photosynthetic and non-photosynthetic reactions, leading to NADP⁺ and N₂ reduction, respectively [187-191]. They are not found in higher eukaryotes. However, through gene fusion events they have been incorporated into CPR and sulphite reductase, and other enzymes [192, 193]. They typically range from 14.5-23 kDa, e.g. cindoxin (16 kDa) from Citrobacter braakii and flavodoxin (FldA) (19.6 kDa) from E. coli [194, 195]. They were first discovered in 1960s in cyanobacteria and Clostrium pasteurianum growing in low iron concentrations, and are thought to have evolved as an alternative to Fe-S proteins [188, 196].

Flavodoxins are separated into two groups, short-chain and long-chain flavodoxins, which are differentiated by a ~20 residue loop (splitting the fifth β-strand of the
Chapter 1: Introduction

central β-sheet) of unknown function [197]. Lopez-Llano et al. showed that removal of this loop from the long-chain flavodoxin from *Anabaena* does not prevent proper folding of the protein, but it does decrease FMN binding affinity [198, 199]. The authors proposed that the differentiating loop of the long-chain flavodoxins may be involved in redox partner recognition. Short-chain flavodoxins typically range from 14-17 kDa, e.g. flavodoxins from *Clostridium beijerincki* and *Desulfovibrio vulgaris* have molecular weights of 15.8 and 16.2 kDa, respectively [200, 201]. Long-chain flavodoxins range from 20-23 kDa, e.g. flavodoxins from *Synechococcus* (sp. PCC 7942) and *Anabaena* (sp. PCC 7119) have molecular weights of 18.7 and 18.9 kDa, respectively [202, 203].

Unlike the FNRs, which can receive one or two electrons, flavodoxins must receive their electrons one by one. Free FMN cannot achieve this as its one-electron reduced form is very unstable [200, 204]. Protein binding alters the redox potentials of the flavin, e.g. the midpoint reduction potential of the SQ/HQ couple is lowered from -172 mV for free FMN to -399 mV at pH 7.0 when bound to flavodoxin from *Clostridium MP* [205]. The midpoint reduction potential of the FMN OX/SQ couple is changed from -238 mV to -92 mV when bound to the same protein. Most flavodoxins function by shuttling between the SQ/HQ states, the latter of which typically has the lower redox potential [206]. The flavodoxin fused to the reductase domain of BM3 functions by cycling OX/SQ, this is due to differences in the loop involved in FMN binding (50s loop), which is one residue shorter in BM3 compared to related flavodoxin proteins [207]. This difference alters the redox potentials of the FMN and is also the reason why, unlike other bacterial flavodoxins which form air stable blue SQs, BM3 stabilises a red anionic SQ [208, 209].

1.2.2.1: The Flavodoxin Fold

The flavodoxin fold is made of a single domain consisting of two α-helical layers sandwiching a five-stranded parallel β-sheet, forming a three-layered single domain protein (Figure 1.11). The FMN is bound on the periphery of the protein, with the isoalloxazine ring shielded from the solvent by a single aromatic residue. Flavodoxins are highly acidic, characterised by two regions containing negatively charged groups, one of which is near the flavin binding region and is likely important for directing interaction with redox partners [210]. Subtle structural changes in the FMN binding
loop region give rise to the differences in properties of the FMN cofactors in the bacterial and eukaryotic CPR enzymes [211].

![Figure 1.11: Structures of bacterial flavodoxins.](image)

The flavodoxin fold consists of a central five stranded β-sheet (red), which is sandwiched by two α-helices on each side (blue). The FMN, shown in stick model, is bound on the protein periphery surrounded by a loop region, which contains most of the residues involved in FMN binding. Flavodoxins shown are from *Anabaena* (sp. 7120) (PDB code; 1FLV) (A) and *Desulfovibrio desulfuricans* (PDB code; 3F6R) (B). Structures were drawn using PyMol.

1.2.2.2: *FMN Binding*

Binding of the FMN cofactor in flavodoxins is similar to the binding of FAD in FNR, and is characterised by three loops, with residues from one loop making hydrogen bonds with the FMN phosphate, and the other two loops involved in interacting with the isoalloxazine ring. The isoalloxazine ring is sandwiched between two aromatic residues, which are found on the 60’s and 90’s loops (Figure 1.12) [203, 212, 213]. In addition, a third aromatic residue is also highly conserved in flavodoxins and lies near the pyrimidine end of the isoalloxazine ring. In the absence of FMN, the aromatic residues move in to fill the void vacated by FMN as demonstrated by flavodoxin from *Anabaena* shown in Figure 1.12, where W57 moves into the flavin binding pocket. A phosphate ion or another anion fills the phosphate binding space [214]. The dimethylbenzene portion of the isoalloxazine ring is normally exposed to the solvent, indicating that electron transfer likely occurs through this portion of the FMN moiety.
Figure 1.12: Superposition of the structures of apo- and holoflavodoxin from *Anabaena*. The isoalloxazine ring of the FMN is sandwiched between two aromatic residues, similar to that seen for FAD binding proteins. FMN binding is also stabilised by hydrogen bonding between the FMN phosphate group and protein backbone residues, and to a lesser extent with the ribityl tail. In the absence of the FMN cofactor, W57 (and to a lesser extent Y94) moves into the vacated space. Structures (PDB codes; 1FTG and 1FLV) were drawn using PyMol, and adapted from [197].

Murray *et al.* demonstrated that isoalloxazine ring binding is dependent on the presence of a phosphate group in the phosphate-binding pocket. They showed that riboflavin does not bind in the absence of inorganic phosphate [215]. Using near UV (NUV) circular dichroism (CD) and NMR spectroscopy, they also showed that W60 movement out of the isoalloxazine binding pocket is facilitated by phosphate binding in *D. vulgaris* flavodoxin. Furthermore, W57 in *Anabaena* flavodoxin forms a hydrogen bond to the phosphate group of FMN. Binding of the phosphate moiety and the subsequent formation of the hydrogen bond likely promotes movement of the W57 out of the isoalloxazine binding pocket [216, 217]. Interestingly, no positively charged residues interact with the phosphate group, and binding is solely through a network of hydrogen bonds with conserved alcholic residues.

1.3: P450 Redox Systems

The monooxygenation reactions catalysed by P450s require a coupled and stepwise supply of electrons, which are usually derived from NAD(P)H and supplied via a
redox partner or partners. The nature of the redox partners varies, and the P450s were historically divided into two major classes according to the different types of electron transfer systems they used. The advent of genome sequencing projects starting in the 90s and continuing today, has led to the discovery of novel P450s in these organisms, e.g. the publication of the M. tuberculosis genome in 1998 by Cole et al. revealed the presence of 20 genes that code for P450 enzymes [218]. Some of the newly identified P450s cannot be grouped into either of the two major classes, and today at least ten classes with subtle differences in redox partners have been described, some of which are shown in Figure 1.13. Some of these redox systems are represented by a single P450 [219]. For instance, CYP176A1 (P450cin) from C. braakii, which utilises a flavodoxin protein instead of a ferredoxin, is the sole member of the class III [220].

Class I P450s include bacterial and mitochondrial enzymes that use a FNR-like enzyme and a ferredoxin protein (Figure 1.13A) [221]. P450cam has served as a model for enzymes belonging to this class. The P450cam redox system includes putidaredoxin reductase (PdR) that is reduced by NADH via hydride ion transfer to its FAD moiety, and putidaredoxin (Pdx), a small Fe-S protein, which can only receive one electron at a time and shuttles electrons between PdR and P450cam. Unlike the class II enzymes, all members of a prokaryotic class I system are soluble proteins and utilise both NADPH and NADH, although most are more selective towards NADH.

The majority of mammalian P450s are called microsomal P450s (e.g. those found in liver) and belong to class II redox systems (Figure 1.13B). These P450s derive their electrons from CPR, a membrane bound diflavin reductase. CPR contains a FAD-binding domain which is homologous to bacterial FNRs that interacts with the pyridine nucleotide coenzyme and a FMN-binding domain that interacts with the P450 [221]. CPR is responsible for the reduction of all microsomal P450s, and receives its electrons from NADPH [183]. Cytochrome b₅ is also capable of delivering the second electron to selective P450s [222]. As mentioned earlier, the mammalian P450s are all membranous enzymes.
Chapter 1: Introduction

Figure 1.13: Schematic organisation of P450 redox systems. (A) In class I bacterial systems, electrons are shuttled from NAD(P)H through a FNR (FdR) to an Fe-S cluster ferredoxin (Fdx) and finally to the P450 heme. (B) Microsomal class II systems shuttle electrons from NADPH through CPR to the P450, both are membrane bound enzymes. Novel systems include XplA (C) and CYP55A1 (D), which do not belong to either of the classical classes. The former contains a P450 component fused to a FMN-binding flavodoxin module (Fldx) and a flavodoxin reductase enzyme (FdR), and the latter does not utilise any redox partners but interacts directly with NADH. Scheme was drawn using Microsoft PowerPoint, and adapted from [117].

The P450 which began the expansion of the two class system was BM3 first identified by Narhi et al. in 1986 [9]. BM3 was unique in that it was the first P450 discovered that was fused to its redox partners forming a completely self-sufficient system only requiring NADPH to function. Since that discovery many other P450s have been identified that are unique in terms of redox partners, facilitating the creation of novel classes (Figure 1.13C-D). Furthermore, there are P450s that can function without the aid of redox partners and instead directly interact with pyridine nucleotide coenzymes, e.g. CYP55A1 (P450nor) from Fusarium oxysporum that receives electrons directly from NADH to catalyse the reduction of nitric oxide (NO) [221, 223]. Moreover, the P450 enzyme system from the archaeon Sulfolobus solfataricus utilises pyruvate as the electron donor [22].

In P450 redox systems, the transfer of electrons normally proceeds through the FAD moiety of a FNR-like reductase enzyme to either a FMN or Fe-S cluster, one electron
at a time. However, there are examples where the FMN can also receive electrons directly from pyridine nucleotide coenzymes in the absence of FAD, e.g. P450 RhF, described in section 1.3.1 below [224].

1.3.1: **P450 Fusion Enzymes**

Although P450s and redox partners are usually expressed independently, in some instances these enzymes have evolved into P450-redox partner and redox partner-redox partner fusion enzymes. Many of these P450-redox partner fusion enzymes form the novel redox systems mentioned above, e.g. MCCYP51FX from *Methylococcus capsulatus* contains CYP51, a sterol 14α-demethylase, fused to a ferredoxin protein [225]. Ultimately, self-sufficient P450 systems have also evolved through the fusion of P450 and CPR genes. The best known example of this type of system is BM3, a self-sufficient and soluble fatty acid hydroxylase [226], catalysing the hydroxylation at ω-1 to ω-3 positions of fatty acids [227]. The fusion of this P450 to its redox partner provides a pathway for efficient and direct electron transfer, and is responsible for the high turnover number of BM3, ~17,000 min⁻¹ with arachidonic acid [9, 228].

Since its discovery, there have been other examples of P450-diflavin reductase enzymes, the most notable of which are the nitric oxide synthases (NOS) [229]. More recently a member of the P450-flavodoxin fusion enzyme family, encoded by the gene *xplA*, has been identified in several strains of *Rhodococci*, and is homologous to the BMP/FMN portion of BM3 (albeit with the order reversed). In addition, P450 RhF from *Rhodococcus* (sp. NCIMB 9784) has an unusual domain arrangement, as it contains a P450 fused to a FMN-containing flavin and Fe-S reductase partner [230]. In this arrangement, the FMN moiety interacts with the pyridine nucleotide coenzyme and accepts the hydride ion.

These fusion enzymes have served as valuable models in the study of P450 and redox partner interactions, with BM3 in particular serving as an excellent model system for studying structure-function relationships in P450s and in analysis of the mechanism of electron transfer. Furthermore, they have been studied extensively in biotechnology as these fusion enzymes are capable of high catalytic turnover, and represent a much simpler system with less redox partners than the more common class I and class II P450 systems.
Chapter 1: Introduction

1.3.2: P450-Redox Partner Interactions

The mechanism of interaction between P450s and their redox partners remains an active area of research. Results from recent studies using covalently linked P450 redox partners supports the proposed mechanism where interaction is driven initially by long range electrostatic interactions, followed by hydrophobic interactions and steric complementarity (Figure 1.14) [231, 232]. The hydrophobic interactions are more important amongst redox partners in some P450 redox systems, whilst in others electrostatic interactions are more important for recognition and binding. It is likely that P450 and redox partner recognition and binding are driven by the same interactions, e.g. Jenkins et al. demonstrated that E. coli flavodoxin binds P450c17 with relatively high affinity ($K_d$ 0.2 μM) at low ionic strength, and that this binding affinity was attenuated at moderate ionic strengths (100-200 mM KCl) [195]. Structural characterisation of a natural P450 fusion enzyme would provide structural details to support these mechanisms. However, structural characterisation of a natural P450 fusion remains elusive.

![Figure 1.14: Structures of natural and covalently linked redox partner fusion enzymes.](image)

(A) Topological view of rat CPR (PDB code; 1AMO), a natural diflavin reductase enzyme consisting of four domains. The FMN-binding domain (blue) shuttles electrons from the FAD- and NADP+-binding domains (light blue) to microsomal P450s. The connecting domain (red) is responsible for correct orientation of the other domains for efficient electron transfer [183]. The FMN, FAD and NADP+ cofactors are shown as yellow, orange and magenta stick models, respectively. (B) Ribbon diagram representation of the dimeric PdR-Pdx fusion enzyme (PDB code; 3LB8) crosslinked via the $K^{409PdR\_E^{72Pdx}}$ pair. The 365 Å² interacting interface is predominantly hydrophobic, and the distance between the PdR FAD isovaloxazine ring and Pdx Fe-S cluster (shown in spheres) is 12 Å [231]. Structures were drawn using PyMol.
Sevrioukova et al. reported the structure of the BMP/FMN domains of BM3 in 1999, shown in Figure 1.15, although the linker region between FMN and heme domains was proteolysed during crystallisation [211]. The FMN domain is positioned at the proximal face of the heme domain in the structure (albeit there being only one FMN domain for each two heme domains in the crystal). FMN-heme interactions are apparently facilitated by long range electrostatic interactions [211]. However, the presence of only a few direct contacts in the 967 Å\(^2\) area of interface indicates interaction between the domains is not strong. Furthermore, the distance between the FMN methyl group and heme iron is 18.4 Å. Studies of redox enzymes by Page et al showed that electrons can travel up to 14 Å between redox centres [233]. It is therefore unlikely that this structure is representative of the catalytically competent electron transfer enzyme. The structure did reveal that the indole ring of W574 might be critical for electron transfer, as it shields the FMN ring from solvent and could act as an electron conduit from FMN to heme [211]. The shallow (proximal face) side of the heme domain contains positive residues that are also possibly critical for redox partner (FMN domain) interactions to enable electron transfer.

Figure 1.15: Structure of the BMP/FMN domains of BM3. The fusion enzyme crystallised with a ratio of two heme domains for each FMN domain. The FMN-binding module is positioned at the proximal face of the heme domain, with the FMN moiety 4 Å away from the protein backbone residue (I385) that precedes the heme-binding loop, and 18.4 Å away from the heme. In contrast to the PdR-Pdx complex, the interacting surfaces are clustered with predominantly negatively charged (FMN domain) and positively charged (BMP) residues. The flexible linker connecting the two domains was proteolysed during crystallisation. The heme and FMN moieties are shown in as sphere models. Structure (PDB code; 1BVY) was drawn using PyMol.
Chapter 1: Introduction

Diflavin reductases and other redox partner fusion enzymes have played an important role in understanding the interactions between redox partners, e.g. their structures demonstrated the role of flexible hinge regions in the modulation of inter-domain electron transfer [234-236]. Furthermore, the structural characterisation of these fusion enzymes and other redox partner complexes has aided in determining the amino acid residues involved in the interacting surfaces, orientation of cofactors and electron delivery pathway, and has provided insight into the electron transfer mechanism [237-240]. In addition, redox partner interaction is known to induce structural changes in both enzymes, not only superficially but also in the protein core [241]. These structural changes likely facilitate electron transfer either by altering redox potentials of the cofactors, or by bringing the cofactors closer. Moser et al. showed that electron transfer rate constants generally decrease exponentially with distance [242]. Therefore, intra-molecular electron transfer is controlled by protein dynamics and thermodynamics [243, 244], and inter-molecular electron transfer is controlled by the rate of complex formation, which must occur to allow efficient inter-protein electron transfer [245].

1.4: The XplA/XplB Redox System

XplA is an unusual cytochrome P450 system comprising a flavodoxin domain fused to the N-terminus of a P450. A partner flavodoxin reductase (encoded by xplB) is required to complete the system. This arrangement suggests that XplA may be the first isolated P450 to have a fused flavodoxin, and might thus represent a completely novel redox system for P450s. The full length enzyme has a molecular weight of 62 kDa. The P450 (XplA-HD) portion of the enzyme shares 29% sequence identity with P450Biol from B. subtilis, and 31% sequence identity with several uncharacterised bacterial P450s. XplA also appears to have evolved to function as a reductase instead of an oxidase, and structural and functional data has been presented to support this theory [106, 246].

XplA was first identified from the soil bacterium R. rhodochrous recovered from areas contaminated with explosive compounds, and was subsequently shown to be able to catalyse the biodegradation of Royal Demolition eXplosive or Research Department Explosive (RDX), chemically known as hexahydro-1,3,5-trinitro-1,3,5-triazene, as a
source of nitrogen \( (K_m = 83.7 \, \mu\text{M} \text{ and } K_d = 58 \, \mu\text{M}) \) \[106, 247\]. Other bacteria and fungi recovered from water, soil and marine environments also contaminated with RDX have since been identified which can also degrade RDX \[248, 249\]. However, these organisms may biotransform RDX utilising an enzyme other than XplA, e.g. the type I nitroreductase enzyme encoded by the \textit{nsfl} gene from \textit{Enterobacter cloacae} \[96-3\] has also been shown to catalyse the breakdown of RDX \[250-253\]. Although many of these organisms may likely possess the \textit{xplA} and \textit{xplB} genes, which are located adjacent to each other in a plasmid-borne gene operon that has already been identified in a host of organisms including other species of \textit{Rhodococcus} and other unrelated bacteria including \textit{Gordonia} \( \text{sp. KTR9} \) and \textit{Williamsia} \( \text{sp. KTR4} \), which have both been shown to contain the \textit{xplA} gene \[254\].

RDX is an environmental xenobiotic compound, and its manufacture, use, storage and disposal has resulted in contamination of the environment and has fuelled the need for a sustainable, low cost method of remediation. It is listed as a “priority pollutant” and as “possible human carcinogen” by the USA Environmental Protection Agency (EPA) \[255, 256\]. Thus, XplA is also of biotechnological importance, and transgenic experiments with plants expressing \textit{xplA} and \textit{xplB} show that they grow faster than wild type plants in soil supplemented with RDX \[106, 257\]. The structure of RDX is shown in Figure 1.16, along with detected major breakdown products, methylenedinitramine (MEDINA) and 4-nitro-2,4-diazabutanal (NDAB). Several mechanisms of RDX breakdown have been proposed, and likely involve reductive denitration followed by an anaerobic- or aerobic-dependent ring cleavage step \[249, 258, 259\]. The removal of RDX in aerobic conditions is two-fold slower than in anaerobic conditions \[257\].
Chapter 1: Introduction

**Figure 1.16: Chemical structures of RDX and major breakdown products.** RDX is a heterocyclic explosive molecule which has been widely used since World War II, leading to widespread contamination of land and groundwater. Its toxicity to all forms of life has made its removal a priority [260-262]. The mechanism of RDX breakdown by XplA is not well understood, and several pathways have been proposed [249, 259]. All of these proceed through reductive denitration followed by ring cleavage of the unstable intermediates, leading to the production of MEDINA or NDAB in anaerobic and aerobic environments, respectively. Structures were drawn using ChemDraw.

The crystal structure of the XplA-HD was recently reported and is shown in Figure 1.17. The structure revealed a constricted active site with a hydrophobic area directly above the heme plane, featuring three clustered methionine residues [246]. The overall structure of XplA-HD resembles the classical prism shaped structure shared by all P450s, with a noticeable difference in the I helix, which is kinked (Figure 1.17A). This disruption is due to the absence of two key conserved residues described in section 1.1.1.3. The aspartate and threonine residues which typically occupy these two positions are replaced by a methionine (M394) and an alanine (A395) (Figure 1.17B). These changes, in addition to the presence of two other methionine residues (M318 and M322), introduce an even more hydrophobic environment in the heme active site compared to other P450s. The structure was solved with an imidazole molecule bound to the heme iron.
XplB is a FAD-binding flavodoxin reductase and is predicted to be the cognate redox partner of XplA. Their genes are adjacent to one another, and functional interactions between these two enzymes were demonstrated by Jackson et al. [106]. The authors also showed that XplB was more selective towards NADPH as the pyridine nucleotide coenzyme, and that NADH did not readily reduce XplB. However, the enzyme remains essentially uncharacterised. XplB has a molecular weight of 48 kDa, and shares 40% amino acid sequence identity with the uncharacterised cindoxin reductase from C. braakii, 31% with FprA from Corynebacterium tuberculostearicum SK141, and ~33% with several putative bacterial FNRs.

1.5: Introduction to the Project

Proteins recruit cofactors to extend the chemistry available within the active site of enzymes. The proteins comprising the P450 redox systems have incorporated heme, FAD, FMN and Fe-S cluster cofactors in order to perform electron transfer reactions ultimately leading to the catalysis of a diverse range of reactions by not only the terminal oxidase P450 enzymes, but also by the flavin containing proteins. The aim of this project is to characterise two enzymes, XplA and XplB, which incorporate three of the four types of redox cofactors that are utilised by proteins in the P450 redox systems.
systems, and which constitute a novel type of P450 redox system of which it is the sole representative characterised to date. Furthermore, the function of XplA is of biotechnological relevance. Determination of the structural and biochemical properties of the XplA/XplB redox system may reveal the catalytic mechanism of RDX degradation by XplA, a P450 which appears to have deviated from the traditional role of P450s as monooxygenases, and into a reductase function. Furthermore, structural characterisation of the intact XplA, in particular, is of great importance in P450 research, due to the elusiveness of a structure of a natural P450-redox partner fusion enzyme. Determination of such a structure would likely provide insight into the mechanisms of P450-redox partner recognition and interaction, and into the process of intra-protein electron transfer. XplA might thus serve as a model for inter-protein transfer, processes of which are reasonably well understood, but hitherto lack structural evidence in the case of the P450s.

BM3 has not only served as model for eukaryotic P450s, but has long been an excellent model for studying structure-function relationships in P450s in general, and for the mechanisms of electron transfer. However, due to its complexity it has often been necessary to separate the fusion enzyme into its individual domains, with the heme/FMN-binding domain representing the simplest model for the electron transfer complex [226]. The discovery of XplA allows this area of P450 research to be conducted on a native and intact enzyme. Therefore the purpose of this project is to undertake expression and purification trials of XplA and XplB to isolate sufficiently pure and stable proteins which would allow detailed structural and functional studies on XplA and XplB using X-ray crystallography and a diverse range of spectroscopic techniques.
Chapter 2: Materials & Methods
Chapter 2: Materials & Methods

2.1: Materials

Expression vectors pET11a, pET15b, competent cells NovaBlue, XL1 Blue, HMS174 (DE3), Rosetta 2 (DE3), BL21 (DE3) and Origami B (DE3) were from Novagen (UK). Growth media, isopropyl β-D-1-thiogalactopyranoside (IPTG), δ-aminolevulinic acid (δ-ALA), and antibiotics ampicillin and carbenicillin were from ForMedium (Hunstanton, UK). Antibiotics kanamycin and rifampicin were from Melford Laboratories (Ipswich, UK). DNA modifying enzymes and buffers, dNTPs, DNA ladders, protein markers, were from New England Biolabs (Hitchin, UK). Chromatography resins Sephacryl S-200, phenyl Sepharose and Q-Sepharose were from Amersham Bioscicenes (Little Chalfont, UK), ceramic hydroxyapatite was from Bio-Rad and Ni-NTA resin was from Qiagen (Crawley, UK). Expression vector pCold™ TF DNA vector was from Takara Bio Inc.

All chemicals and reagents used for experimental procedures were purchased from Sigma-Aldrich (Poole, UK) or Fisher Scientific (Loughborough, UK). RDX was purchased from Vitas-M Laboratory, Ltd (Moscow, Russia).

Table 2.1: Buffer compositions. Potassium phosphate buffers were made to the correct pH by mixing 1 M solutions of mono- and dibasic potassium phosphate. Tris- and HEPES-based buffers were adjusted to the desired pH with hydrochloric acid.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Constituents</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>50 mM potassium phosphate, 250 mM KCl, 10% glycerol (pH 8.0)</td>
</tr>
<tr>
<td>B</td>
<td>10 mM HEPES, 150 mM NaCl (pH 7.5)</td>
</tr>
<tr>
<td>C</td>
<td>50 mM HEPES (pH 7.5)</td>
</tr>
<tr>
<td>D</td>
<td>25 mM potassium phosphate (pH 6.5)</td>
</tr>
<tr>
<td>E</td>
<td>50 mM HEPES, 150 mM NaCl (pH 7.2)</td>
</tr>
<tr>
<td>Redox</td>
<td>100 mM potassium phosphate, 150 mM NaCl, 10% glycerol (pH 7.0)</td>
</tr>
<tr>
<td>TE</td>
<td>50 mM Tris, 1 mM EDTA (pH 7.2)</td>
</tr>
</tbody>
</table>
2.2: Experimental Methods

2.2.1: Restriction Digestions

The gene of interest was excised from the vector by using the appropriate restriction enzymes. Typically, the digestion mixture contained 3 µL of plasmid DNA (300 ng), 4 µL of 10x buffer optimal for the restriction enzymes used, 4 µL of 10x bovine serum albumin (BSA), 1.5 µL of each restriction enzyme required (30 U) and made up to 40 µL with sterile distilled deionised (dH$_2$O), and then incubated at 37 °C for 4 h. In the event of sub-optimal digestion, an additional 1.5 µL of restriction enzyme (30 U) was added after 2 h to improve digestion efficiency. Digestions to isolate vectors were performed as for the insert, but the digestion mixture incorporated 1.5 µL of calf intestinal alkaline phosphatase (CIP, 15 U) to remove the 5'-phosphate groups from vector DNA to prevent self-ligation. Alternatively, 1 µL Antarctic Phosphatase (5 U) and 10x reaction buffer (NEB) were added and incubated for 15 min, following digestion.

The resulting digests were run on a 0.8% (w/v) agarose gel to separate vector from insert. The agarose gel was prepared as follows, 200 mg of agarose was dissolved in 25 mL of 1x Tris Acetate EDTA buffer (TAE) (40 mM Tris, 20 mM acetic acid, 1 mM EDTA (pH 7.6)), to which 3 µL of ethidium bromide (10 mg/mL stock) was added and mixed. The agarose solution was then poured into a casting tray with combs, and allowed to set. The samples were prepared by adding 2 µL of 6x bromophenol loading dye and applied onto the gel, and subjected to electrophoresis at 80 V for 1 h in 1x TAE buffer using a Fisherbrand MH-675 horizontal gel unit. 12 µL of 2.4x diluted 2-log DNA ladder (New England Biolabs) (100 µL ladder, 100 µL dH$_2$O and 40 µL 6x loading dye) was run alongside the samples to confirm fragment size, and unless stated otherwise the same diluted DNA ladder was used in all DNA gels. Following electrophoresis, the position of the restriction digested-derived DNA and gene fragments were visualised using a GENE FLASH Syngene Bio Imaging system and the gel was photographed using a Computar H6Z0812 lens attached to a Sony UP-895MD printer. Fragments of the appropriate size were excised from the gel and purified from the agarose using the QIAquick gel extraction kit (Qiagen), according to the manufacturer’s instructions.
Alternatively, restriction digestions were done on a small scale when confirming the presence of a gene of interest. Typically the digestion mixture contained 2 μL of plasmid DNA (200 ng), 1.5 μL of 10x BSA, 0.5 μL of each restriction enzyme (10 U), 1.5 μL of 10x buffer suitable for every restriction enzyme in the digest, and made up to 15 μL with dH2O.

2.2.2: Transformations

*E. coli* competent cells were used to inoculate 5 mL Luria Bertani (LB) medium supplemented with antibiotics selective for that strain, and allowed to grow at 37 °C with shaking overnight (O/N) to generate the starter culture. The starter culture was used to inoculate 5 mL of LB medium supplemented with appropriate antibiotics and grown at 37 °C until OD600 ~0.5, at which point 1 mL aliquots were taken per transformation. The inoculant volume was typically 1% (v/v) of the new culture volume, i.e. 50 μL of the starter culture was used to inoculate 5 mL. The cells were subsequently centrifuged at 13,000 rpm for 1 min and 4 °C using a Beckman Coulter Microfuge 22R centrifuge. The supernatant was discarded and the pellet resuspended in 0.5 mL of sterile, pre-chilled 50 mM CaCl2. The resuspended cells were left on ice for 30 min. 1 μL of plasmid DNA (typically 0.1 μg DNA) was then added, gently mixed and left on ice for a further 2 min. The cells were ‘heat shocked’ at 42 °C for 30 s and the samples returned to ice for 2 min. Then, 80 μL of pre-warmed SOC medium was added and the transformants were incubated for a further 30 min and 37 °C with shaking. The cells were centrifuged as above and the supernatant poured off and resuspended in residual supernatant prior to plating out onto pre-warmed LB agar selective plates and incubating at 37 °C O/N. For fast growing strains, like HMS174 (DE3), 150 μL were plated directly onto agar selective plates and incubated at 37 °C O/N.

Successful transformants from non-expression host strains were selected and used to inoculate 5 mL LB medium supplemented with antibiotics selective towards the plasmid to generate a starter culture as above. The starter culture was used to inoculate either 5 mL or 50 mL of LB media (for mini- and midi-prep, respectively) and cells were grown at 37 °C until OD600 ~0.5. The culture was centrifuged at 4,500 rpm for 10 min and 4 °C using a Heraeus Labofuge 400R centrifuge. Plasmid DNA
was prepared by mini- or midi-prep spin kit (Qiagen, West Sussex, UK), following the manufacturer's guide. The presence of the gene of interest was verified by restriction digestion with appropriate restriction enzymes and visualisation of the insert on a DNA agarose gel. DNA sequencing of clones (Eurofins MWG Operon) confirmed correct gene sequences. Subsequently, the plasmid DNA was transformed into *E. coli* expression strains for heterologous recombinant protein expression. Alternatively, the plasmid DNA was re-transformed into a non-expression host strain to maintain a glycerol stock by adding 200 µL of sterile 80% (v/v) glycerol to 800 µL of cell culture and storing at -80 °C. Glycerol stocks of transformants from expression host strains were also kept frozen, as above.

### 2.2.3: Cloning

Vector and insert DNA for ligation reactions were cut using the same restriction enzymes to ensure complementary ends, as described above. Ligation reactions were done in 0.2 mL PCR tubes. The ligation mixture typically contained 1 µL of vector (50 ng), 2 µL of insert, 1 µL of 10x ligase buffer, 1 µL of hexamine cobalt chloride (HCC) (1 mM), 1 µL of T4 ligase (0.4 Weiss U), and was made up to 10 µL with dH₂O and incubated at 4 °C O/N. Various stoichiometric ratios of vector to insert from 2:1 to 1:5 were tested to maximise the chance of successful ligation. Alternatively, the reaction was done using the Quick Ligase kit (NEB), according to the manufacturer's protocol.

The resulting construct was transformed into non-expression *E. coli* cells as previously described. Successful cloning was established by screening for transformants on antibiotic selective LB agar plates. Plasmid DNA was prepared by mini- or midi-prep and the resulting DNA preparations were verified by restriction digest. DNA sequencing of clones (Eurofins MWG Operon) confirmed correct gene sequences. Positive clones were subsequently transformed into one or more *E. coli* expression host strains.
Chapter 2: Materials & Methods

2.2.4: Site-directed Mutagenesis

The site-directed mutagenesis polymerase chain reaction (PCR) experiments were done using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene), following the manufacturer’s protocol [263]. Complementary oligonucleotide primers that anneal to the gene sequence to be modified were designed using the QuikChange Primer Design tool to include the desired mutation or mutations and were synthesised by Eurofins MWG Operon [263]. The PCR parameters were adjusted for the size of the plasmid following the manufacturer’s protocol. Following the PCR experiment, the template DNA was destroyed using the restriction enzyme DpnI, which only hydrolysates methylated DNA. PCR amplified mutated DNA is not methylated. Plasmids from successful transformants of the mutated DNA were sequenced to confirm the presence of the desired mutation and digested with the relevant restriction enzymes to excise the gene, which could be visualised on a DNA gel.

2.2.5: Expression Trials

Expression trials were conducted on a small scale in 50 mL cultures. Expression host cells carrying the desired construct were grown at 37 °C until an OD$_{600}$ of 0.6-0.8 was reached. At this point, a 1 mL aliquot of culture was taken, and gene expression induced by supplementing with 1 mM IPTG and incubation at 37 °C, taking 1 mL aliquots at 1, 2, 4 and finally 24 h after induction. The OD$_{600}$ reading was noted at each collection time point. A control culture with no IPTG induction was also grown concurrently to test basal expression levels of the gene, and samples were taken as with the induced sample. The 1 mL samples collected were centrifuged at 13,000 rpm for 10 min and 4 °C on a benchtop microfuge, the supernatant was discarded and the pellet was resuspended in 200 µL 50 mM Tris, 1 mM EDTA (pH 7.2) buffer (TE). 12 µL of the normalised cell resuspension was heated with 6 µL of 3x SDS sample buffer at 95 °C for 3 min and analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Normalised loading volumes for SDS-PAGE analysis were calculated using the worksheet provided in the pET System Manual (Novagen). 10 µL of protein marker, broad range (2-212 kDa) (New England Biolabs) was run
alongside the samples to confirm band size, and unless stated otherwise the same protein marker was used in all SDS-PAGE gels.

To determine the solubility of the expressed protein, expression was performed as above, but 1 mL aliquots were not taken and instead the entire 50 mL culture was harvested by centrifugation at 4,500 rpm for 30 min and 4 °C after the optimum induction period, as determined in the previous experiment. The cells were lysed using BugBuster (BB) (Novagen), using 5 mL of 1x BB per mg of wet cell pellet. The resuspended sample was supplemented with lysozyme (10 μg/mL), DNase (10 μg/mL), protease MINI inhibitor cocktail (Roche), and left on a shaking platform for 20 min at room temperature (R/T). The bacterial cell lysate was cleared of cellular debris by centrifugation at 13,000 rpm for 30 min and 4 °C. The soluble and insoluble fractions were analysed by SDS-PAGE. The *E. coli* strain which produced the largest quantity of soluble protein, if any, was subsequently chosen to further the expression trials. The trials were done with variations in cell strain, expression temperature, and IPTG concentration to optimise the protocol to be used in later large scale expression. Trials also included the addition of riboflavin (5 ng/mL) and δ-ALA (90 μM) for flavo- and hemoproteins, respectively.

To determine the medium that produces the highest yield of soluble protein, the expression was scaled up to 500 mL of LB, 2x yeast tryptone (YT) and terrific broth (TB) media. Gene expression was induced by supplementing the cultures with the optimum concentration of IPTG as determined in previous trials. 250 mL of the cultures were harvested by centrifugation at 6,500 rpm for 12 min and 4 °C in a Beckman Coulter J-26 XP centrifuge (JLA 16.250 rotor) after 2 h of expression and the rest after 24 h expression. Cell pellets were resuspended in 20 mL 50 mM potassium phosphate (KPi), 250 mM KCl, 10% glycerol (pH 8.0) (buffer A), supplemented with lysozyme (10 μg/mL), DNase (10 μg/mL), protease inhibitor cocktail, and lysed by sonication with 20 s bursts, at 50 s intervals and 50% amplitude for 30 min on a Sonics Vibra-cell sonicator. Cell debris was cleared by centrifugation at 20,000 rpm for 30 min and 4 °C in a Beckman Coulter J-26 XP centrifuge (JA 25.50 rotor). The supernatant was loaded directly onto a Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed with five column volumes of buffer A, followed by five column volumes of buffer A with 20 mM imidazole. The protein was eluted using
Chapter 2: Materials & Methods

a one step gradient with 300 mM imidazole in buffer A. 1 mL fractions were collected at each step and analysed by SDS-PAGE.

2.2.6: Expression & Purification of XplA

The expression plasmid pET15b(xplA) was transformed into HMS174 (DE3) (Novagen) cells. Conditions used for cell growth and protein expression were 37 °C in 1 L cultures of LB medium, until OD600 reached ~0.6-0.8. At this point the temperature was lowered to 20 °C and cells were allowed to equilibrate prior to induction of xplA expression by addition of 100 µM IPTG and cell growth for a further 24 h. Cells were harvested by centrifugation (6,000 rpm, 15 min, 4 °C) using a Beckman JLA8.1 rotor. The cell pellet was washed in ice cold TE buffer, and stored frozen at -20 °C. Typically a total of 24 L of culture was grown and harvested at a time, and the pellets combined.

The frozen pellet was thawed in buffer A with 10 mM imidazole supplemented with protease inhibitor cocktail (Roche), lysozyme (10 µg/mL), and DNase (10 µg/mL). Cells were lysed by sonication with 15 s bursts, at 50 s intervals and 40% amplitude for 30 min using a Bandelin SONOPULS sonicator. The temperature was kept at <10 °C. Cell debris was removed by centrifugation at 20,000 rpm for 30 min and 4 °C. The supernatant was loaded directly onto a Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed with five column volumes of buffer A, followed by five column volumes of buffer A with 20 mM imidazole. The protein was eluted using a one step gradient with 100 mM imidazole in buffer A. Alternatively, the column was washed in 5 mM histidine in buffer A prior to elution using 100 mM histidine. The partially purified XplA protein was concentrated in an Amicon ultrafiltration device with a 30 kDa molecular weight cut off (MWCO) and subsequently diluted with TE buffer. It was then dialysed against 2 litres of the same buffer with four changes of buffer. Dialysed protein was filtered using a 0.22 µM filter and concentrated using Vivaspin 20 mL concentrators (30 kDa MWCO, Generon) to ~500 µL final volume. Concentrated protein was loaded onto a Superdex 200 10/300 GL (GE Healthcare) gel filtration column (24 mL, flow rate 0.5 mL/min) equilibrated with 10 mM HEPES, 150 mM NaCl (pH 7.5) (buffer B), collecting 1 mL fractions. Fractions containing XplA were analysed for purity with SDS-PAGE and by their A418/A280 ratio. The purest
fractions were pooled and concentrated using Vivaspin 20 mL concentrators to ~1 mL, and dialysed against buffer B plus 50% glycerol prior to storage at -80 °C.

2.2.7: Expression & Purification of XplA-HD

The expression plasmid pET15b(xpla-HD) encoding the XplA heme domain was transformed into HMS174 (DE3) cells. Conditions used for cell growth and protein expression were identical to the intact XplA.

The frozen pellet was thawed in 50 mM HEPES (pH 7.5) (buffer C) supplemented with protease inhibitor cocktail, lysozyme (10 µg/mL), and DNase (10 µg/mL). Cells were lysed and cellular debris cleared as described for the intact XplA. The supernatant was loaded directly onto a DEAE ion exchange column (Amersham) equilibrated with buffer C. The column was washed with five column volumes of buffer C. The protein was eluted using a linear gradient of 0-500 mM KCl in buffer C. The eluent was collected in 5 mL fractions. Fractions containing XplA-HD were identified visually and were pooled and concentrated in an Amicon with a 30 kDa MWCO and subsequently diluted with 25 mM KPi (pH 6.5) (buffer D). The partially purified XplA-HD was loaded onto a hydroxyapatite column equilibrated with buffer D. The column was washed with five column volumes of buffer D, and XplA-HD was eluted using a linear gradient of phosphate in the concentration range 25 mM to 500 mM. The eluent was collected in 2 mL fractions. Fractions containing XplA-HD were analysed for purity using their A₄₁₈/A₂₈₀ ratio, and those with a ratio greater 1:1 were pooled, filtered using a 0.22 µM filter and concentrated using Vivaspin 20 mL concentrators to ~500 µL final volume. Concentrated protein was purified further using size-exclusion chromatography as described for the intact XplA. Fractions containing XplA-HD were analysed for purity by their A₄₁₈/A₂₈₀ ratio. The purest fractions with a A₄₁₈/A₂₈₀ ratio greater than 1.4:1 were pooled and concentrated using Vivaspin 20 mL concentrators to ~1 mL, and dialysed against buffer B plus 50% glycerol prior to storage at -80 °C.

Alternatively, XplA-HD was purified using Ni-NTA affinity and Superdex 200 10/300 GL size-exclusion chromatographic columns as described for the intact XplA.
Chapter 2: Materials & Methods

2.2.8: Expression & Purification of XplB

The expression plasmid pET15b(xplB) was transformed into HMS174 (DE3) cells. Conditions used for cell growth and protein expression were 37 °C in 0.5 L cultures of TB medium, until OD$_{600}$ reached ~0.6-0.8. At this point induction of xplB expression was done as with xplA, and the cells harvested as previously described in section 2.2.6. Typically a total of 12 L of culture was grown and harvested at a time, and the pellets combined.

The frozen pellet was lysed as described for XplA in section 2.2.6, and the supernatant loaded directly onto a Ni-NTA column equilibrated with buffer A with 10 mM imidazole. The column was washed with five column volumes of buffer A, followed by five column volumes of buffer A with 30 mM imidazole. The protein was eluted using a one step gradient with 250 mM imidazole in buffer A. The partially purified XplB protein was concentrated in an Amicon with a 30 kDa MWCO and subsequently diluted with buffer D, and loaded on a hydroxyapatite column equilibrated with buffer D. The column was washed with five column volumes of buffer D, and the protein eluted using a linear gradient of the phosphate buffer in the concentration range 25 mM to 500 mM collecting 2 mL fractions. UV/Visible spectroscopy was used to identify fractions containing XplB, and the purity of the protein was analysed using SDS-PAGE. The purest fractions were pooled and concentrated using Vivaspin 20 mL concentrators with a 30 kDa MWCO to ~500 µL final volume. Concentrated protein was loaded onto a Superdex 200 10/300 GL gel filtration column as described for XplA, collecting 1 mL fractions. Fractions containing XplB were identified using UV/Visible spectroscopy and were analysed for purity with SDS-PAGE. The purest fractions were pooled and concentrated using Vivaspin 20 mL concentrators to ~1 mL, and dialysed against buffer B plus 50% glycerol prior to storage at -80 °C.

2.2.9: Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

SDS-PAGE is an analytical strategy used to separate molecules according to their size [264, 265]. Adjustment of the acrylamide percentage (w/v) controls the porosity of the gel and optimises it for proteins of a particular molecular weight range. SDS is a negatively charged detergent that is used to solubilise, denature and linearise proteins and give them a uniform charge-to-size ratio, so that they can be separated
solely on the basis of their size. SDS, which binds polypeptides in a constant weight ratio of 1.4 g SDS/g of polypeptide, has a high negative charge that overwhelms any charge that the protein may have, imparting all proteins with a relatively equal negative charge [266].

SDS-PAGE gels were cast between glass plates (BioRad, Hertfordshire, UK). A 10% (w/v) acrylamide resolving gel (5 mL) was prepared as follows; 1.9 mL of dH2O, 1.3 mL of 1.5 M Tris-HCl (pH 8.8), 1.7 mL of acrylamide (30% (w/v) acrylamide: 0.8% (w/v) bis-acrylamide stock solution (37.5:1)) and 50 μL of 10% (w/v) SDS. Polymerisation was initiated by addition of 2 μL of N, N, N’, N’-tetramethylethylenediamine (TEMED) and 50 μL of 10% (w/v) ammonium persulphate (APS), immediately before loading into the gel plates. Approximately 3-4 cm was left between the top of the resolving gel solution and the top of the gel plates. Isopropanol (1 mL) was layered on top of the resolving gel to smooth the surface and remove bubbles. Polymerisation was allowed to proceed for at least 15 min, and once set the isopropanol was removed. In order to ensure that all proteins entered the resolving gel at the same time, a 5% (w/v) acrylamide stacking gel was used on top of the main resolving gel. The stacking gel (1 mL) was prepared as follows; 0.68 mL of dH2O, 0.13 mL of 1.0 M Tris-HCl (pH 6.8), 0.17 mL of acrylamide stock solution and 10 μL of 10% (w/v) SDS. Polymerisation was initiated as above with 10 μL of 10% (w/v) APS and 1 μL of TEMED, and the solution added immediately on top of the polymerised resolving gel. A 10 well comb was placed into the stacking gel and allowed to set as above.

The gel was placed into a vertical electrophoresis mini-PROTEAN 3 system (BioRad, Hertfordshire, UK). Both the inner and outer reservoirs were filled with running buffer (25 mM Tris base, 192 mM glycine (pH 8.3), 0.1% (w/v) SDS) prepared from a 10x concentrated stock. Protein samples (typically 10 μL) were denatured by mixing with 2x SDS sample buffer and heating at 95 °C for 3 min. Protein marker, broad range (2-212 kDa) was loaded (10 μL) into one of the lanes for reference and the gel was run at a constant voltage of 180 V until the electrophoretic front had reached the bottom of the gel. Bromophenol blue added to the sample buffer is used as a tracking dye to indicate the progress of the electrophoresis run. As the dye is small and negatively charged it moves under the influence of the electrophoretic field, faster than the proteins in the sample.
Chapter 2: Materials & Methods

To detect proteins in the SDS-PAGE gels, the gels were stained with Coomassie stain as follows; each gel was rinsed with water and stained for 10 min with gentle shaking in staining solution (40% (v/v) ethanol, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie Brilliant Blue (G-250)) at R/T until the protein bands became clearly visible. The gel was then de-stained by washing in dH₂O until a clear background was obtained.

2.2.10: Preparation of FMN-free XplA

In order to generate FMN-free XplA, the enzyme was diluted in 50 mM K₂HPO₄, 1 M (NH₄)₂SO₄, 1 mM EDTA (pH 7.0) (equilibration buffer) and immobilised on a phenyl-Sepharose column. The flavin was then removed using the method of Lederer [267]. Briefly, bound protein was washed with 50 mM K₂HPO₄, 1 M (NH₄)₂SO₄, 1 mM EDTA (pH 4.5) and the flow-through monitored for fluorescence on an Eclipse fluorescence spectrophotometer. Quartz cells of 1 cm pathlength were used and excitation was at 450 nm for the FMN (with emission collected between 500-600 nm). The column was washed extensively until no fluorescence was detected. The column was re-equilibrated with equilibration buffer and the FMN-free XplA was subsequently eluted using a linear gradient consisting of 25 mM KPi, 1 mM EDTA (pH 7.0) (elution buffer) and dH₂O. Eluted apoprotein was concentrated using a 20 mL Vivaspin Concentrator with a 30 kDa MWCO, dialysed against buffer B plus 50% glycerol and stored at -80 °C.

2.2.11: Flavin Quantification

The flavin content of XplA was quantified using the method of Aliverti [268]. Following gel filtration, a sample from the purest fractions was diluted to 1 mL with 10 mM HEPES (pH 7.5). The concentration of this protein solution was determined by absorbance difference spectrum, as described in 2.2.14. The protein was recovered, transferred to a 1.5 mL Eppendorf and wrapped in foil. The flavin was subsequently removed by heat denaturation at 100 °C for 10 min. The solution was then cooled on ice and the precipitated protein was removed by centrifugation at 13,000 rpm for 10 min at 4 °C. The fluorescence emission spectrum of the supernatant between 480-600 nm was recorded using a Varian Cary Eclipse Fluorescence spectrophotometer with an excitation wavelength of 450 nm and
excitation and emission slit openings at 10 nm. The emission intensity at 524 nm of the recovered flavin was corrected for buffer emission and compared to the emission intensities of known concentrations of free FMN. The range of linearity of the instrument was determined by measuring the emission spectra of solutions containing increasing concentrations of FMN in the same buffer. Alternatively, the concentration of FMN in the supernatant was determined using an extinction coefficient of $\varepsilon_{446} = 12.2 \text{ mM}^{-1} \text{ cm}^{-1}$ for the oxidised flavin absorption spectrum [269].

2.2.12: Flavin Reconstitution

Purified XplB was loaded onto a Ni-NTA column equilibrated with buffer A. The column was washed with three column volumes of buffer A, followed by 10 column volumes of 2 M guanidinium chloride and 2 M KBr in buffer A to elute the flavin. The column was then washed with 10 column volumes of buffer A. Three column volumes of 10 mM FAD (Sigma-Aldrich) were cycled through the column 10 times, or O/N. The FAD was left on the column for >2 h, then washed off with five column volumes of buffer A. XplB was eluted using buffer A containing 250 mM imidazole.

XplA (0.5-1 mM) was reconstituted with FMN by incubation with 10 mM FMN in 50 mM HEPES, 150 mM NaCl (pH 7.2) buffer O/N. Excess FMN was removed by passing the enzyme solution down a PD-10 column (Bio-Rad).

2.2.13: Reconstitution of the XplA/XplB Redox System

Reconstitution of the XplA/B redox system was done in an anaerobic glove box (Belle Technology, Weymouth UK) to maintain oxygen levels at <2 ppm. XplB-dependent XplA activity was determined by monitoring RDX-dependent NADPH oxidation. Concentrated protein stock solutions (~500 μM) were made anaerobic by passing them down a PD-10 column equilibrated with anaerobic buffer. Concentrated solutions of FAD (1 mM), NADPH (100 mM) and RDX (25 mM) were made up in the glove box using the same buffer. The reaction mixture contained 5 μM XplA, 25 μM FAD, 25 μM XplB and 100 μM RDX. The reaction was initiated by addition of 200 μM NADPH. Control experiments were performed in the absence of either XplB or FAD. The concentration of XplB was determined using an extinction coefficient of $\varepsilon_{280} =$
2.2.14: UV/Visible Spectroscopy

UV/Visible absorption spectra were recorded on a Cary UV-50 Bio UV/Visible scanning spectrophotometer (Varian, UK) using 1 cm pathlength quartz cells. Unless otherwise stated, spectra were recorded using ~5-10 μM of enzyme in 50 mM HEPES, 150 mM NaCl (pH 7.2) (buffer E). The concentration of the XplA was initially determined using the method of Omura and Sato [17], using an extinction coefficient of $\varepsilon_{450} = 91 \text{ mM}^{-1} \text{ cm}^{-1}$ in the reduced CO-bound minus reduced absorption difference spectrum. CO binding to XplA was performed by reduction with a few grains of solid sodium dithionite, followed by slow bubbling of CO into the reaction mixture. NO binding was performed by passing a few bubbles of gas into an oxidised enzyme solution. Spectral data for the XplA complex with β-mercaptoethanol (BME) were collected in 50 mM KPi at pH values from 6-8.

Optical titrations to determine $K_d$ values for XplA ligands were carried out at 20 °C with ligands solubilised in appropriate solvents. Spectra were recorded (250-800 nm) after each addition of the ligand. RDX, trinitrobenzene (TNB) and polycyclic azole drugs were solubilised in dimethyl sulfoxide (DMSO). Dithiothreitol (DTT) and BME were solubilised in buffer E. 4-phenylimidazole (4-PIM) solutions were in 50% ethanol/buffer E mixtures. Unless otherwise stated, other ligands used were prepared in buffer E, and binding titrations were done in the same buffer.

Following completion of titrations, difference spectra were generated by subtraction of the ligand-free spectrum from those generated at each point in the titration. The $K_d$ values were determined by fitting the data for the ligand-induced absorption change (normally peak minus trough data, using the same wavelength pair for each difference spectrum in a particular titration) versus ligand concentration to either a standard (Michaelis-Menten) hyperbolic function (Equation 2.1) or to a quadratic
function (Equation 2.2) for tight binding ligands, using Origin software (OriginLab, Northampton, MA).

**Equation 2.1:**

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

In Equation 2.1, \( A_{obs} \) is the observed absorbance change at ligand concentration \([L]\), \( A_{max} \) is the absorbance change at ligand saturation and \( K_d \) is the dissociation constant for the enzyme-ligand complex.

**Equation 2.2:**

\[
A_{obs} = \left(\frac{A_{max}}{2[E]}\right) \left(([L] + [E] + K_d) - \left(([L] + [E] + K_d)^2 - (4[L][E])\right)^{0.5}\right)
\]

In Equation 2.2, \( A_{obs} \) is the observed absorbance change at ligand concentration \([L]\), \( A_{max} \) is the absorbance change at ligand saturation, \([E]\) is the enzyme concentration and \( K_d \) is the dissociation constant for the enzyme-ligand complex [271].

In studies to determine the \( K_d \) value for TNB, the apparent \( K_d \) value for the competitive binding of RDX (which induces a large optical shift on binding XplA) was determined at three different concentrations of TNB (0.25, 0.5 and 1 mM) and data fitted using Equation 2.1. The \( K_d \) for TNB was then determined using Equation 2.3.

**Equation 2.3:**

\[
K_i = \frac{K_d \times [TNB]}{K_{d\text{ app}} - K_d}
\]

In Equation 2.3, \( K_d \) is the dissociation constant for RDX determined in absence of TNB, \([TNB]\) is the concentration of TNB used in the relevant titration with RDX, \( K_{d\text{ app}} \) is the apparent \( K_d \) for RDX at the given \([TNB]\), and \( K_i \) is the \( K_d \) value for TNB. The \( K_d \) reported for TNB is then the average of the three \( K_i \) values determined at the three different TNB concentrations.
Chapter 2: Materials & Methods

2.2.15: Fluorescence Spectroscopy

To determine the binding affinity between XplA and FMN, a 250 nM FMN solution was prepared in 10 mM HEPES (pH 7.5). This FMN solution was titrated against 500 µM FMN-free XplA, with FMN delivered in 0.5 µL aliquots. One minute was allowed to elapse between each titration to allow equilibration. The fluorescence emission spectrum between 480-600 nm was measured as before, with an excitation wavelength of 450 nm and excitation/emission slits openings at 20 nm. The titration was stopped once additions of FMN-free XplA caused no further decrease in fluorescence.

To account for non-specific fluorescence changes and dilution factors, a second titration experiment was done with a 250 nM FMN solution titrated with buffer to the same final volume. To determine the FMN binding constant, the intensity change at 524 nm following each titration addition was corrected for buffer contribution and subtracted from the intensity of the free FMN, and then plotted against the concentration of apoprotein.

2.2.16: Pyridine Hemochromogen Assay

The pyridine hemochromogen method was used to quantify XplA heme and to determine an extinction coefficient at the XplA Soret maximum. Analysis was done according to the method of Berry and Trumpower [272]. UV/Visible absorbance spectrum was recorded for a 10 µM XplA solution (1 mL), 0.5 mL of the protein solution was removed and 0.5 mL of pyridine stock solution (200 mM NaOH, 0.8 mM K₃Fe(CN)₆, 40% (v/v) pyridine) was added, and the spectrum of the oxidised hemichrome taken. The hemichrome complex was reduced using sodium dithionite, and the spectrum of the hemochrome complex taken. Spectra were recorded until no further changes occurred. The concentration of XplA was calculated using the difference in absorbance between the spectra for the oxidised and reduced samples, using a difference extinction coefficient of ε₅₅₅ = 23.98 mM⁻¹ cm⁻¹. This concentration was multiplied by two and applied to the Beer-Lambert law (Equation 2.4) to calculate the extinction coefficient for the Soret peak. Multiple samples were analysed to ensure consistency of determined extinction coefficients for the Soret peak.
Equation 2.4:

\[ A_{obs} = \varepsilon cl \]

In Equation 2.4, \( A_{obs} \) is the absorbance of the Soret, \( \varepsilon \) is the extinction coefficient, \( c \) is the enzyme concentration in mM as determined from the pyridine hemochromogen method, and \( l \) is the pathlength of the sample cuvette in cm.

2.2.17: Redox Potentiometry

To determine the XplA heme iron \( \text{Fe}^{3+}/\text{Fe}^{2+} \), FMN OX/SQ, and FMN SQ/HQ midpoint reduction potentials, redox titrations were performed in an anaerobic glove-box (Belle Technology) under a nitrogen atmosphere with \( O_2 \) levels maintained at less than 2 ppm. All solutions were deoxygenated by sparging with nitrogen gas.

A concentrated sample of XplA (~1 mM) was applied to a PD-10 desalting column in the anaerobic box, pre-equilibrated with anaerobic 100 mM KPi, 150 mM NaCl, 10% (v/v) glycerol (pH 7.0) (redox buffer), to remove traces of \( O_2 \). A 5 µM protein solution was titrated electrochemically according to the method of Dutton using a concentrated sodium dithionite solution as reductant [273]. Mediators were added to facilitate electrical communication between enzyme and electrode, prior to titration. Typically, 2 µM phenazine methosulfate (+80 mV), 7 µM 2-hydroxy-1,4-naphthoquinone (-145 mV), 0.3 µM methyl viologen (-430 mV), and 1 µM benzyl viologen (-311 mV) were included to mediate in the range between +100 to -480 mV. Midpoint potentials of the mediators are indicated. Dithionite was delivered in ~0.2 µL aliquots and at least 10 minutes were allowed to elapse between each addition to allow equilibration to occur. Spectra (250-800 nm) were recorded using a Cary UV-50 Bio UV/Visible scanning spectrophotometer. The electrochemical potential of the solution was measured using a Hanna pH 211 meter coupled to a Calomel electrode (ThermoRussell Ltd.) at 25 °C. The electrode was calibrated using the \( \text{Fe}^{3+}/\text{Fe}^{2+} \) EDTA couple as a standard (+108 mV). A factor of +244 mV was used to correct relative to the standard hydrogen electrode.

To determine the heme iron reduction potential from the redox titration data collected, spectral data at wavelengths diagnostic for the optical (Soret) transition
Chapter 2: Materials & Methods

between oxidised (Fe$^{3+}$) and reduced (Fe$^{2+}$) forms of P450 heme iron were plotted versus the applied potential, and the data fitted using the Nernst function with Origin software. To determine the FMN reduction potentials, spectral data at 630 nm (a wavelength that reports on optical transitions between OX, 1-electron reduced (SQ) and 2-electron reduced (HQ) forms of the FMN without significant influence by heme absorption) were plotted versus the applied potential, and the data fitted using a two electron Nernst function, as described in previous studies [274, 275].

2.2.18: Stopped-flow Spectroscopy

Stopped-flow absorption measurements were made using an Applied Photophysics SX18 MVR stopped-flow spectrophotometer in an anaerobic glove box (Belle Technology) to maintain oxygen levels at <2 ppm. Multiple wavelength data were collected using a photodiode array (PDA) detector and XSCAN software. XplA-CO binding rate constants were measured at 25 °C by monitoring formation of the ferrous-CO complex at 448 nm in buffer C. All experiments were carried out with anaerobic buffer saturated by extensive bubbling with CO gas. Reactions were initiated by mixing a solution of dithionite-reduced enzyme (10 μM) with a solution containing various concentrations of CO (23-139 μM from a 975 μM saturated stock solution) [276]. Data were fitted using a single exponential function using Spectrakinetics software (Applied Photophysics), and the observed reaction rate constants ($k_{obs}$) were plotted against the relevant [CO] using Origin software and fitted to a linear function to obtain the 2nd order CO binding rate constant.

To analyse the formation of the XplA oxyferrous complex, the XplA-HD (~500 μM) was first mixed with excess dithionite under anaerobic conditions. A sample was taken, sealed and analysed by UV/Visible spectroscopy to confirm full reduction. The reduced sample was then passed down a PD-10 column in the glove box to remove excess dithionite. The eluted sample was diluted to ~10 μM and checked as before to confirm that the heme was still fully reduced. The reduced sample was then mixed in the stopped-flow instrument with buffer C containing varying concentrations of O$_2$ (20-100 μM from a 250 μM air saturated stock solution) [277]. Entire spectral accumulation allowed identification of the spectral changes involved in formation of the oxyferrous complex of XplA-HD, and absorbance changes at 435 nm (where there
is a large increase in absorption for the oxyferrous complex) were recorded against time in single wavelength mode. Data were fitted using a single exponential function and the 2nd order $O_2$ binding rate constant was obtained as described for CO binding. XplA-O$_2$ binding rate constants were measured at 4 °C.

![Figure 2.1: Derivation of a second order rate constant from stopped-flow kinetics. Illustration describes the derivation of the rate constants of ligand (L) binding, $k_{on}$, and ligand dissociation, $k_{off}$, from the gradient and y-axis intercept, respectively, of the linear plot of individual rate constants ($k_{obs}$) versus (L).](image)

2.2.19: Electron Paramagnetic Resonance Spectroscopy

EPR spectra for the ligand-free and ligand-bound forms of XplA were recorded on a Bruker ER-300D series electromagnet with microwave source interfaced with a Bruker EMX control unit and fitted with an ESR-9 liquid helium flow cryostat (Oxford Instruments), and a dual-mode microwave cavity from Bruker (ER-4116DM). Spectra were recorded at 10 K with a microwave power of 2.08 mW and modulation amplitude of 1 mT. Oxidised XplA (300 μM) was prepared in 50 mM HEPES (pH 7.2) plus 150 mM NaCl (200 μL). Ligand concentrations used were at least 10x their $K_d$ value. EPR spectra were collected by Dr Stephen Rigby (University of Manchester).
**Chapter 2: Materials & Methods**

### 2.2.20: Magnetic Circular Dichroism Spectroscopy

Magnetic circular dichroism (MCD) spectra were recorded using JASCO J/810 and J/730 dichrographs in the NUV/Visible and near infra-red regions (NIR) regions, respectively, using an Oxford Instruments superconducting solenoid with a 25 mm ambient bore to generate a magnetic field of 6 T. A 0.1 cm pathlength quartz cuvette was used to record NIR spectra with sample concentrations the same as those used for EPR data collection. UV/Visible spectra were recorded for XplA with 50 mM HEPES in 2H2O (pH* 7.0) as buffer (where pH* is the apparent pH measured in 2H2O using a standard glass pH electrode). MCD spectra were recorded by Dr Hazel Girvan (University of Manchester) and Dr Myles Cheesman (School of Chemistry, University of East Anglia).

### 2.2.21: Circular Dichroism Spectroscopy

CD spectra were recorded on a JASCO J-810 spectrapolarimeter, at 20 °C, and in 50 mM KPi, 300 mM KCl buffer (pH 7.5). Far UV CD spectra (FUV) (180-260 nm) were collected in quartz cuvettes with pathlengths ranging from 0.1-0.5 mm using a scan rate of 10 nm min⁻¹. Protein concentrations used were typically 2 μM (200 μL), although both parameters were adjusted depending on the cuvette used and the signal intensity of the sample. NUV/Visible CD spectra (260-600 nm) were collected in 2 mm pathlength quartz cuvettes with a protein concentration of 20 μM (1 mL) using a scan rate of 20 nm min⁻¹. CD spectra for buffer used in experiments were also recorded and subtracted from spectra of the samples. Spectra were recorded in triplets and averaged. The secondary structural content was analysed using the K2D3 program [278].

### 2.2.22: Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) was used to determine the binding constant for XplB and FAD. XplB was prepared by dialysing in degassed buffer E, and made up to a final concentration of 25 μM using the dialysed buffer. FAD was made up to 500 μM using the same buffer used to dialyse XplB to ensure identical solutions. The same buffer was also used to wash the instrument. The concentration of XplB was
determined as described in section (2.2.13). Experiments were performed at 25 °C using a VP-ITC micro calorimeter (MicroCal). 8 μL aliquots of FAD were titrated into the protein at a sample injection speed of 5.7 μL/s. Injections were made at 300 s intervals. A control experiment was performed by injecting FAD into buffer. The heat of dilution was negligible and was subtracted from the final data prior to analysis. Data were analysed using ITC-Expert software (MicroCal) and Origin 7.

2.2.23: **Multiangle Laser Light Scattering**

Protein samples used for multiangle laser light scattering (MALLS) were in the concentration range of 1-5 μM in 50 mM KPi, 300 mM KCl (pH 7.5) buffer. Buffers were prepared, filtered and degassed on the day of the experiments. A 500 μL sample was run on a size-exclusion chromatography column (S-200) using a flow rate of 0.71 mL/min. The eluent passed through a DAWN-EOS MALS spectrometer (Wyatt Technology corp. Santa Barbara CA, USA) detector with an Optilab rEX refractometer (Wyatt Technologies) and a quasi-elastic light scattering (QELS) detector (Wyatt) to measure the refractive index and the hydrodynamic radius (Rh) values, respectively. The average molecular mass (Mr) was calculated using ASTRA v5.21 (Wyatt Technology Corporation) and were derived using Zimm fitting [136, 173]. MALLS experiments were conducted by Mrs. Marjorie Howard at the Biomolecular Interactions Facility in the Faculty of Life Sciences, University of Manchester.

2.2.24: **Thermofluor Assay**

Experiments were performed using a Bio-Rad CFX96 real-time system C100 thermal cycler, using 96-well RT-PCR plates (Bio-Rad Laboratories, Hercules, CA, USA). Measurements were performed using wavelengths for excitation and emission of 490 and 575 nm, respectively. Unfolding curves were generated using a temperature gradient from 15 to 95 °C, performing a fluorescence measurement after every 0.2 °C increase after a 5 s delay for signal stabilisation. Fluorescence is measured and plotted against temperature, enabling determination of the melting temperature (Tm). In the Thermofluor software, Tm is typically determined by assigning the minimum of the corresponding inverted first derivative curve. All experiments were performed at
least three times, and the reported $T_m$ values are based on the mean values determined from the peaks of the derivatives of the experimental data.

In a typical buffer screening experiment, 1–2 μL of a concentrated protein solution was mixed with 4.5 μL of 500x SYPRO orange dye (90x final), 12.5 μL of 100 mM buffer from the JBS Solubility Kit (Jena Bioscience) (50 mM final) from Table 2.2, directly into the plate wells, and the mixture diluted with dH$_2$O to a final volume of 25 μL. The buffer screening experiments were used to identify the buffer condition which gave the optimal $T_m$ for the protein of interest, and this was selected for further screening using the additive kit, the second component of the JBS Solubility Kit shown in Table 2.3. In a typical additive screening experiment, 1–2 μL of a concentrated protein solution was mixed with 4.5 μL of 500x SYPRO orange dye, and then 12.5 μL of the selected buffer and 5 μL of the additive (Table 2.3) were mixed directly into the plate wells and diluted with dH$_2$O to a final volume of 25 μL. Control samples were run concurrently and contained the protein without additives.

A range of protein concentrations were tested for each protein to determine the concentration which gave a good signal to noise ratio, which was typically between 0.5 to 5 mg/mL. Protein stock solutions were made up in dH$_2$O, and SYPRO orange dye was diluted using dH$_2$O. The additive screen was repeated with ligand-bound enzyme. Final ligand concentrations used were 10x their $K_d$ value, and were pre-bound in the protein stock solution.
### Table 2.2: JBS Solubility Kit part A; Buffer kit (100 mM).

<table>
<thead>
<tr>
<th>No.</th>
<th>Buffer</th>
<th>pH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Glycine</td>
<td>3.0</td>
</tr>
<tr>
<td>2</td>
<td>Citric acid</td>
<td>3.2</td>
</tr>
<tr>
<td>3</td>
<td>PIPPS</td>
<td>3.7</td>
</tr>
<tr>
<td>4</td>
<td>Citric acid</td>
<td>4.0</td>
</tr>
<tr>
<td>5</td>
<td>Sodium acetate</td>
<td>4.5</td>
</tr>
<tr>
<td>6</td>
<td>Sodium/potassium phosphate</td>
<td>5.0</td>
</tr>
<tr>
<td>7</td>
<td>Sodium citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>8</td>
<td>Sodium/potassium phosphate</td>
<td>6.0</td>
</tr>
<tr>
<td>9</td>
<td>Bis-Tris</td>
<td>6.0</td>
</tr>
<tr>
<td>10</td>
<td>MES</td>
<td>6.2</td>
</tr>
<tr>
<td>11</td>
<td>ADA</td>
<td>6.5</td>
</tr>
<tr>
<td>12</td>
<td>Bis-Tris propane</td>
<td>6.5</td>
</tr>
<tr>
<td>13</td>
<td>Ammonium acetate</td>
<td>7.0</td>
</tr>
<tr>
<td>14</td>
<td>MOPS</td>
<td>7.0</td>
</tr>
<tr>
<td>15</td>
<td>Sodium/potassium phosphate</td>
<td>7.0</td>
</tr>
<tr>
<td>16</td>
<td>HEPES</td>
<td>7.5</td>
</tr>
<tr>
<td>17</td>
<td>Tris</td>
<td>7.5</td>
</tr>
<tr>
<td>18</td>
<td>EPPS</td>
<td>8.0</td>
</tr>
<tr>
<td>19</td>
<td>Imidazole</td>
<td>8.0</td>
</tr>
<tr>
<td>20</td>
<td>Bicine</td>
<td>8.5</td>
</tr>
<tr>
<td>21</td>
<td>Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>22</td>
<td>CHES</td>
<td>9.0</td>
</tr>
<tr>
<td>23</td>
<td>CHES</td>
<td>9.5</td>
</tr>
<tr>
<td>24</td>
<td>CHAPS</td>
<td>10.0</td>
</tr>
</tbody>
</table>

### Table 2.3: JBS Solubility Kit part B; Additive kit.

<table>
<thead>
<tr>
<th>No.</th>
<th>Additive</th>
<th>Concentrated stock solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sodium chloride</td>
<td>80 mM</td>
</tr>
<tr>
<td>2</td>
<td>Sodium chloride</td>
<td>200 mM</td>
</tr>
<tr>
<td>3</td>
<td>Sodium chloride</td>
<td>400 mM</td>
</tr>
<tr>
<td>4</td>
<td>Glycerol</td>
<td>20%</td>
</tr>
<tr>
<td>5</td>
<td>Glycerol</td>
<td>40%</td>
</tr>
<tr>
<td>6</td>
<td>CHAPS</td>
<td>8 mM</td>
</tr>
<tr>
<td>7</td>
<td>Octyl glucopyranoside</td>
<td>0.4%</td>
</tr>
<tr>
<td>8</td>
<td>Octyl glucopyranoside</td>
<td>4%</td>
</tr>
<tr>
<td>9</td>
<td>Dodecyl maltoside</td>
<td>0.4%</td>
</tr>
<tr>
<td>10</td>
<td>Dodecyl maltoside</td>
<td>4%</td>
</tr>
<tr>
<td>11</td>
<td>BMF</td>
<td>40 mM</td>
</tr>
<tr>
<td>12</td>
<td>DTT</td>
<td>4 mM</td>
</tr>
<tr>
<td>13</td>
<td>DTT</td>
<td>20 mM</td>
</tr>
<tr>
<td>14</td>
<td>TCEP</td>
<td>120 mM</td>
</tr>
</tbody>
</table>
Chapter 2: Materials & Methods

2.2.25: Crystallisation Trials

Initial crystallisation trials were conducted on a small scale using 100 nL drops using a Mosquito pipetting robot (Molecular Dimensions, New Market). Protein concentrations from 5 to 50 mg/mL were tested, and ligand concentrations equal to the protein concentration plus an amount equal to 10x its dissociation constant to assure saturation were used. Protein samples were tested against common crystallisation conditions (Molecular Dimensions) using 96-well crystallography trays. Trials for the intact XplA were also supplemented with 1 mM FMN in the mother liquor and protein solutions. In some experiments the substrate RDX (1 mM) was also supplemented in the mother liquor. Trials were conducted at 4 °C and at 20 °C.

Protein conditions which were conducive for crystal growth were selected for further optimisation. Buffer conditions which were conducive to crystal growth were reproduced and used in scaled up crystallogenesis experiments using drop sizes from 0.2–2 μL. Alternatively, in the event of multiple hits from a commercial crystallisation screen, the whole screen was re-used. Scaled up commercial screens were conducted in 96-well trays, and manual trays using 2 μL drops and reproduced buffers were done in 24-well plates.

2.2.26: XplA-HD Crystallisation & Structural Analysis

The XplA-HD was crystallised using pure protein concentrated to 25 mg/mL in 10 mM HEPES (pH 7.5). Diffraction quality crystals were obtained using the sitting drop method by mixing equal volumes of protein solution (plus 1 mM RDX) with mother liquor from a number of conditions using a Molecular Dimensions Morpheus screening kit (400 nL total volume), and by incubating at 4 °C and at R/T. Crystals appeared within 2-3 days when incubated at 4 °C, and within one day with incubation at R/T. A single crystal obtained using 50 mM MgCl₂, 0.1 M imidazole/MES (pH 6.5), 30% PEG 550MME-PEG 20K was flash-cooled in liquid nitrogen and diffraction data collected at beamline IO4 of the Diamond synchrotron (Didcot, UK). Data were reduced and scaled with the X-ray Detector Software suite (XDS) [279]. Diffraction data were observed to 2.3 Å and the crystal belonged to the 141 space group with cell parameters a=b=136.12 Å, c=75.13 Å. The structure was solved using molecular
replacement with the available XplA heme domain structure (PDB code; 2WIY). Positional and B-factor refinement was carried out using Refmac5 from the CCP4-suite and data collection and final refinement statistics are given in Table 4.1 in Chapter 4 (4.6.2) [280-282].
Chapter 2: Materials & Methods
Chapter 3: Spectroscopic Characterisation of XplA
Chapter 3: Results I

3.1: Introduction

The cytochrome P450-flavodoxin fusion enzyme XplA is unique amongst the P450-redox partner fusion enzymes, where a bacterial flavodoxin protein is fused to the N-terminus of a P450. It is so far the only known P450 identified with this domain arrangement. Furthermore, together with its cognate redox partner XplB, a FAD binding NADPH-dependent flavodoxin reductase, represent a novel class of P450 redox system, of which XplA is the only member. XplA confers in *R. rhodochrous* the ability to metabolise the xenobiotic compound RDX as a source of nitrogen. The enzyme has been successfully expressed in transgenic plants which were subsequently able to grow faster in soil supplemented with RDX, demonstrating its bioremediation potential [257].

Our focus was drawn to XplA in light of a long-standing interest in P450-redox partner fusion enzymes [27], but also as a consequence of some unusual spectral features reported for this enzyme, including a red shifted Soret maximum and absence of a defined absorbance shoulder in the 450-460 nm region that would be typical for a flavin-binding protein [15,21]. Although details have been published on the apparent function of XplA and its potential bioremediation applications [106, 257, 283, 284], its biochemical and biophysical properties have not been well characterised.

In order to gain a better understanding of the biochemistry of XplA, the enzyme was heterologously expressed in *E. coli*, purified and characterised using a variety of spectroscopic techniques including UV/Visible, EPR and MCD spectroscopy. Furthermore, the XplA/XplB redox system presents an opportunity to study the mechanism of electron transfer from redox partners to P450 and of the nature of P450/redox partner interactions. Elucidation of the ligand binding and thermodynamic properties of XplA may reveal new details on this novel P450, including identifying other potential substrates, and on its role in *R. rhodochrous* and other bacteria in which it is also present.

3.2: Expression & Purification of XplA

The *xplA* gene was codon-optimised and synthesised by GeneService and cloned into the expression vector pET15b using *Ndel* and *BamHI* restriction sites prior to the
start of the project. The pET15b vector carries an N-terminal His$_6$-tag, which enables recombinant proteins to be purified on a nickel affinity chromatography column, and gene expression is under the control of a T7 lac promoter. The vector also carries the gene for ampicillin resistance to allow for selection of cells carrying the plasmid. The entire sequence is displayed in Figure A1 of the addendum.

This construct (pET15b(xplA)) was transformed into tetracycline-resistant XL1 Blue competent cells, and was used primarily to maintain a glycerol stock to be stored at -80 °C from which fresh plasmid could be prepared when required. The construct was also transformed into three E. coli strains containing the DE3 lysogen, namely BL21 (DE3), Rosetta 2 (DE3) and HMS174 (DE3), and expression trials were conducted in those strains. The DE3 lysogen contains the T7 RNA polymerase gene needed for recombinant protein expression. This gene is under the control of the lacUV5 promoter, and expression is induced by the addition of IPTG.

To determine whether the protein was expressed in these strains, 50 mL cultures were grown and gene expression induced using IPTG. Samples from all three cultures were taken at various time points (1, 2, 4 and 24 h after induction). These samples were run on a SDS-PAGE gel, and an overexpressed protein of ~62 kDa could be seen from all three samples (Figure 3.1). This band was confirmed to be XplA by protein gel identification using mass spectrometry. The expression of XplA was significantly improved compared to basal expression following induction with IPTG.
Chapter 3: Results

Figure 3.1: Time-lapse expression of *R. rhodochrous* xplA in *E. coli*. 10% (w/v) SDS-PAGE gel showing the expression profile of XplA in HMS174 (DE3) cells with a band corresponding to the expected size of XplA (~62 kDa) indicated by the marked arrow. The profile shows (1 mL) samples taken from several time points from cells induced with 100 µM IPTG, and from cells not induced with IPTG as the control. The gel shows from left to right; molecular weight marker (1), time = 0 (2), 1 h control (3), 1 h induced (4), 2 h control (5), 2 h induced (6), 4 h control (7), 4 h induced (8), 24 h control (9) and 24 h induced (10).

To establish the solubility of the expressed protein, samples were taken from each of the three strains and the cells lysed using BB as described in Materials & Methods (2.2.3). The soluble fraction was separated from the inclusion bodies by centrifugation, and analysed using SDS-PAGE. Soluble protein was recovered from all three samples. The expression trials showed that HMS174 (DE3) cells produced the largest quantity of soluble protein, as determined by comparison of the band intensities in the SDS-PAGE gels after normalisation of the loading based on optical densities of the cell cultures. Subsequently, HMS174 (DE3) cells were selected for further optimisation. This involved varying growth conditions such as the temperature, IPTG concentration, induction period, growth media and addition of supplements such as riboflavin and δ-ALA. Ultimately, it was determined that HMS174 (DE3) cells grown in LB medium and induced with 100 µM IPTG at 20 °C for 24 h produced the largest quantity of soluble protein. Addition of riboflavin or δ-ALA, or both, did not improve the yield of soluble XplA. Furthermore, the addition of these two additives did not alter heme or FMN incorporation by XplA.

XplA was purified using two chromatographic steps using a Ni-NTA affinity column and a size-exclusion column (Sephadex S-200) to purify from the soluble cell lysate.
and to polish the purified protein, respectively. SDS-PAGE analysis of Ni-NTA purified XplA showed that highly purified protein was recovered (Figure 3.2), and pure XplA achieved an $A_{418}/A_{280}$ ratio of 1:1. Approximately 29 mg of homogenous protein was purified from 12 L of culture (~2.4 mg/L).

![Image of SDS-PAGE gel](attachment:SDS-PAGE.png)

**Figure 3.2: Ni-NTA affinity chromatography purification of XplA.** 10% (w/v) SDS-PAGE gel showing from left to right; molecular weight marker (1), soluble fraction of cell lysate (2), eluate from a 10 mM imidazole wash (3) and eluate from a 100 mM imidazole wash with a major band corresponding to the expected size of XplA (~62 kDa) (4). Expression conditions were; 100 µM IPTG induction at 20 °C for 24 h in HMS174 (DE3) cells.

Occasionally, indole hydroxylation and indigo production was observed in the *E. coli* cells following *xplA* gene induction by IPTG. The cell pellet was of a dark green-blue (almost black) colour after centrifugation, which was isolated in the inclusion bodies following separation from the red-coloured heme (Figure 3.3). The conditions that encouraged indole hydroxylation were not reproducible, and it is possibly linked to the activity of the tryptophanase enzyme, which catalyses the breakdown of tryptophan into indole and pyruvate [285]. The expression of tryptophanase in *E. coli* cells has been shown to be inhibited by catabolite repression, and is sensitive to aeration, glucose concentration and the presence of other metabolites and amino acids [286-288]. All of these conditions are affected by the length of the logarithmic growth phase of the cells, and of the optical density of the cells during induction with IPTG. Therefore under the right conditions, tryptophanase will be expressed leading to high levels of the indole in the growth medium, which can subsequently be hydroxylated by the expressed XplA using native redox partners.
3.3: UV/Visible Spectroscopy

3.3.1: Spectroscopic Properties of XplA

Initial spectroscopic analysis of the purified protein showed that XplA exhibited a typical P450 spectrum with distinct α and β peaks at 570 nm and 541 nm, respectively (Figure 3.4A; red). However, the absorption maximum corresponding to the heme Soret peak was at 421 nm, which is red shifted compared to most low-spin (LS) P450s, which are normally centred at ~418-419 nm, e.g. CYP130 at 418 nm and CYP121 at 416.5 nm [289, 290]. Absorption maxima at these wavelengths are indicative of the resting ferric LS Cys\(^{-}/\)H\(_2\)O ligated enzyme. In addition, the α peak appears more as a shoulder of the β peak, rather than being a distinctive peak of its own. This is likely due to a ligand bound to the heme other than a water molecule, and possibly imidazole despite dialysis steps to remove it. However, there are examples of P450s that display a UV/Visible spectrum with absorption maximum at ~421 nm when in the ferric LS hexa-coordinated form, such as CYP124 from *M. tuberculosis*, which has an absorption maximum at 421 nm [291]. The more prominent δ peak at ~364 nm when compared to the oxidised enzyme, which
Chapter 3: Results I

typically demonstrates a flat and broader feature, is also suggestive of the ligation of a nitrogenous ligand. More extensive dialysis of the protein shifts the XplA Soret peak to 417 nm with concomitant flattening of the δ peak and an increase in the intensity of the α peak (Figure 3.4A; black). The spectral changes at these wavelengths are consistent with the removal of imidazole.

Interestingly, despite the dialysis steps to remove imidazole it was occasionally observed with some preparations of XplA that the absorption maximum remained at a wavelength longer than 417 nm. Subsequently, it was determined that Tris is also able to ligate directly to the heme iron to induce a type II shift similar to that observed for imidazole ligation, where the Soret peak shifts from 417 nm to a longer wavelength due to direct binding of a ligand to the heme iron. The absorption maximum was at 417 nm when the sample was analysed in KPi or HEPES buffers. To confirm Tris ligation to XplA, an optical titration of XplA with Tris was performed. The spectral changes induced by Tris binding are consistent with a type II ligand binding to a P450, and the spectral changes mirrored those of imidazole binding to XplA, albeit the binding was a lot weaker than imidazole with a $K_d$ clearly in the mM range (Figure 3.4B).

Figure 3.4: Spectroscopic characterisation of purified XplA. (A) UV/Visible absorption spectra of XplA (4 µM) following Ni-NTA affinity and size-exclusion chromatography purification. The positions of all absorption peaks arising from the heme are shown and labelled (α-δ). The Soret peak (γ) of XplA is initially positioned at 421 nm (red), and after extensive dialysis shifts to 417 nm (black), with concomitant flattening of the δ peak and increase in absorption at the α peak. (B) XplA (5 µM) in 50 mM HEPES, 150 mM NaCl (pH 7.2) buffer (black) and in 50 mM Tris, 1 mM EDTA (pH 7.2) buffer (red). The Soret peak is at 423 nm in buffer containing Tris, with other spectral features similar to imidazole ligation in the Q-band and δ regions, suggesting ligation of Tris to the XplA heme. The enzyme in HEPES buffer displays spectral features expected of a P450 in the resting ferric LS Cys/H$_2$O ligated state.
Following the resolution of the issues with isolating XplA with water ligated as the 6th axial heme ligand, the spectroscopic properties of XplA were further characterised. XplA shows a characteristic P450 spectrum for the substrate-free ferric enzyme, predominantly in the LS state as characterised by a lack of a HS shoulder at ~390 nm (Figure 3.5; black). The Soret peak is at 417 nm with the α and β peaks at 566 and 540 nm, respectively. On reduction to the ferrous enzyme with sodium dithionite the α and β peaks merge into a single broad peak at ~542 nm and the Soret peak blue shifts to 408 nm (Figure 3.5; blue). Addition of CO to the ferrous enzyme induces a shift of the Soret peak to 446.5 nm, corresponding to the reduced and CO-bound heme iron with the cysteine ligand in the thiolate state, which is commonly referred to as the P450 complex (Figure 3.5; red) [20, 292]. A small shoulder at ~420 nm is also observed as a result of the protonation of the thiolate ligand to form the thiol-ligated CO complex known as the P420 complex [13, 292]. Furthermore, the single broad peak in the Q-band region narrows and shifts to ~548 nm. Also shown in Figure 3.5 (in magenta) is XplA bound to NO to form a Fe(III)NO adduct, where the Soret peak shifts to 431 nm with more distinct α and β peaks compared to the oxidised enzyme at 573 and 543 nm, respectively. The peak at ~362 nm is due to excess NO.
Figure 3.5: Spectroscopic characterisation of ferric, ferrous, ferrous/CO-bound and ferric/NO-bound XplA. The UV/Visible absorption spectra of XplA (5 µM) in the resting ferric LS Cys/H₂O ligated (black), dithionite-reduced ferrous (blue), ferrous/CO-bound (red) and ferric/NO-bound (magenta) states are shown. Soret absorption maxima are located at 417, 408, 446.5 and 431 nm, respectively. Distinct α and β bands are visible in the ferric spectrum at 566 and 540 nm. These merge into a single broad peak at ~542 nm in the ferrous spectrum, which shifts to ~548 nm in the CO-bound spectrum. A small shoulder near 420 nm is visible in the CO-bound species due to protonation of the cysteine thiolate forming a small population of the P420 complex. On NO binding to the ferric enzyme, the Soret peak shifts to 431 nm and more prominent α and β peaks emerge, when compared to the oxidised enzyme, at 573 and 543 nm, respectively. All four species were characterised aerobically.

Examination of the spectrum of the oxidised XplA suggested that FMN might be present in a sub-stoichiometric quantity by comparison with the heme. This conclusion arises as a consequence of the lack of any strong absorption shoulder in the ~450-480 nm region that would be expected if an oxidised FMN cofactor was stoichiometrically bound (analysis is subject to Chapter 5; 5.6.1).

3.3.2: Ligand Binding Studies

3.3.2.1: Type I Ligands

Type I ligands bind to the active sites of P450s in a position different to the one occupied by the axial 6th heme ligand, normally water. Therefore, they do not ligate
directly to the heme iron. However, the binding of type I ligands does displace the water ligand, leading to a transition from LS to HS in the heme iron and a consequent positive shift in heme iron potential, and usually starting the P450 catalytic cycle. RDX is the only ligand hitherto reported to induce a type I spectral shift in XplA [106]. RDX, being a synthetic compound, may not be the natural substrate for XplA, and as with many bacterial P450s its natural substrate is not yet known. Jackson et al. reported that XplA bound RDX with a $K_d$ value of 58 $\mu$M, and suggested that RDX was its natural substrate [106]. However, on inspection of the published spectrum of the titration of XplA with RDX, non-standard spectral changes and other peculiarities in the spectrum were noticed. The reported Soret peak was at ~421 nm with a very prominent $\delta$ band. This is consistent with partial ligation of a nitrogenous ligand to the heme iron as described earlier. Furthermore, the spectral changes were not characteristic of substrate binding to a P450, with a trough at 390 nm that appears to transform into a peak as the titration proceeds. It is likely that the reported binding affinity may be an overestimation, and an optical titration of XplA with RDX was thus performed to establish the true $K_d$ value of RDX binding.

The results of the titration shows spectral changes more consistent with substrate binding to a P450, where the Soret peak starts at 417 nm and gradually blue shifts to 396 nm as the titration proceeds, and has an isosbestic point with LS XplA at ~408 nm (Figure 3.6A). The wavelength of maximum absorption is characteristic of a heme iron in the HS state. There is also a diminution of the $\alpha$ peak and a feature appearing at ~651 nm which is a charge transfer (CT) species between iron and cysteine sulphur that is typical of substrate-bound P450s [94, 293, 294]. A plot of the difference in changes in absorption at 420 and 390 nm against the concentration of RDX titrated produced a binding curve that was fitted using Equation 2.1 to produce an apparent $K_d$ of 7.47 ± 0.12 $\mu$M (Figure 3.6B). This shows that XplA binds RDX significantly tighter than previously reported by Jackson et al., with a near 10-fold decrease in the apparent $K_d$ value [106], and confirms that the originally reported UV/Visible absorption spectrum of XplA titrated with RDX and subsequent derivation of the $K_d$ from the titration was misrepresented, possibly due to the presence of another ligand.
Figure 3.6: Optical titration of XplA with RDX. (A) UV/Visible absorption spectra of XplA (3.5 μM) titrated with RDX. Arrows indicate directions of absorbance change during the titration. Aliquots (0.1 μL) of RDX (25 mM stock) were added to a final concentration of 100 μM. The Soret peak in the initial state is at 417 nm and blue shifts to 396 nm, typical of a P450 transitioning from a ferric LS state to a HS state. The inset shows the initial (black) and final (blue) spectra from the titration. (B) Binding curve from the titration generated by subtracting the absorption changes at 390 nm from those at 420 nm. The data were fitted using Equation 2.1 generating an apparent $K_d$ of $7.47 \pm 0.12$ μM. The inset shows the difference spectra from which the wavelengths of maximum absorption changes were determined and used to generate the binding curve.
In light of the observation that *E. coli* cell cultures expressing *xplA* could hydroxylate indole, we investigated the possibility of indole being a substrate. To accomplish this, an optical titration of XplA with indole was performed. However, no spectral changes were induced which might have shown that indole binding was associated with a HS heme iron shift (Figure 3.7). Although interactions between indole and XplA were not detected spectroscopically, it is reasonable to assume that they do so *in vivo*. Indole, being a hydrophobic molecule, will likely enter the active site of XplA approaching close to the heme iron and initiate catalytic activity. Subsequently, indigo is produced as a consequence of dimerisation of the oxidised indole molecules [295]. The efficiency of indigo production is a concentration dependent process, in relation to the amounts of oxidised indole products formed. As described earlier, the process also likely involves the tryptophanase enzyme responsible for indole production from tryptophan [285].

![Figure 3.7: XplA interactions with indole.](image)

**Figure 3.7: XplA interactions with indole.** UV/Visible absorption spectra of XplA (3.0 μM) titrated with indole. The spectra show minimal spectral changes following indole addition. Only the initial (black) and final (red) spectra from the titration are shown, corresponding to the enzyme in the ferric ligand-free state and following addition of 100 mM indole. The hydrophobic indole molecule will likely enter the XplA active site and approach near the heme iron, although the spectra shows it is unlikely to bind to the heme either as a type I substrate or type II (heme-coordinating) inhibitor. However, low level indole oxidation activity catalysed by XplA, likely explains the formation of indigo in *E. coli* expression cells.
3.3.2.2: Displacement of CO by RDX

During anaerobic CO trapping assays with XplA and P450 redox partners (*E. coli* FldR or XplB and NADPH) (Chapter 5; 5.5.5), it was observed that upon the addition of RDX to the reaction mixture, the enzyme shifted to the substrate-bound HS state and not to ~446.5 nm as expected with P450 Fe(II)CO complex formation. Further, upon the addition of RDX to XplA bound to CO, RDX was able to oxidise the ferrous heme and displace CO from the P450 complex. To confirm this observation a series of CO displacement experiments were performed. XplA in CO-saturated buffer was reduced with 50 µM dithionite to form the P450 complex (Figure 3.8; red), and 1 mM RDX was then added to fully displace the bound CO and form HS heme with the Soret peak shifting to 396 nm after 2 min (Figure 3.8; blue). A further addition of 2 mM dithionite completely reduced the RDX after 10 min (both free RDX and the RDX bound to the XplA), returning the XplA heme to the ferric LS state (Figure 3.8; magenta).

![Figure 3.8: RDX-dependent oxidation of CO-bound ferrous heme of XplA.](image)

**Figure 3.8:** RDX-dependent oxidation of CO-bound ferrous heme of XplA. UV/Visible absorption spectra of XplA (3 µM, in CO-saturated buffer) in the ferric ligand-free (black), and ferrous CO-bound (red) states following the addition of 50 µM dithionite. The P450 complex fully collapses 2 min after the addition of 1 mM RDX and a type I-induced Soret peak appears at 396 nm with a HS CT band at ~651 nm (blue). On addition of 2 mM dithionite, the RDX is fully reduced after 10 min and the Soret peak shifts to 422 nm (magenta).
Chapter 3: Results I

At higher concentrations of dithionite (100 μM dithionite, 1 mM RDX), conversion to HS heme was incomplete due to partial reduction of RDX by excess dithionite. These results show that dithionite is capable of directly reducing the nitro groups of the RDX molecule. Moreover, although XplA returned to the ferric LS state, the Soret peak did not return to 417 or 408 nm corresponding to the ferric and ferrous states, but to 422 nm. In addition, the Q-band region is similar to that observed for the oxidised enzyme, albeit with a slightly more pronounced β peak. This suggests possible ligation of the heme iron by a dithionite-dependent degradation product of RDX or a product of RDX catalysis by XplA. It is reasonable to assume that ligation of this breakdown molecule prevents the reduction of the heme iron by dithionite possibly due to a very low reduction potential of this ligand-bound species, thus the Soret peak remains at 422 nm, and not at 408 nm as expected of a heme iron in the ferrous LS state.

The bound CO could not be displaced by imidazole despite its apparently high affinity when compared to RDX (estimated by its difficulty in removal). This would seem to suggest that the CO is not being directly displaced by RDX but, as the enzyme exist in an equilibrium between bound and unbound states, the Fe(II) rapidly reduces RDX when the CO dissociates, restoring heme to the Fe(III) state, which does not bind CO. Subsequent binding of a RDX breakdown product then traps XplA in a ligand-bound ferric state rather than in the water-ligated resting state.

3.3.2.3: Type II Ligands

As briefly mentioned earlier, type II ligands bind directly to the heme iron of P450s. In addition, like the type I ligands they displace the bound water ligand but approach closer to the heme. Binding of type II ligands typically shifts the Soret peak to longer wavelengths compared to the resting enzyme. This wavelength is dependent on the ligand field strength as described by the spectrochemical series [296, 297]. The high field ligands in the series, such as cyanide and CO, induce more extensive Soret red shifts due to a higher population of the LS heme species.

Cyanide is a potent inhibitor of human mitochondrial cytochrome c oxidase, but can also bind to P450s. Similar to CO, which forms complexes with P450s and famously haemoglobin, cyanide is a potent P450 inhibitor and is toxic to humans. Cyanide is one of the strongest ligands in the spectrochemical series, and induces a large
splitting $\Delta$ of the heme iron $d$-orbitals. Therefore, P450 complexes with cyanide invariably form LS heme. Optical titrations of XplA with sodium cyanide induced spectral changes consistent with direct heme ligation of a strong field ligand (Figure 3.9A). The Soret peak red shifts from 417 nm to 436 nm, indicating a further transition to the LS state, and a much more prominent $\delta$ peak emerges at 367 nm. In addition, there is a merging of the $\alpha$ and $\beta$ peaks in the Q-band region to form a single peak at $\sim$559 nm. The dissociation constant for cyanide was determined by fitting induced absorption changes versus cyanide concentration using Equation 2.1, yielding a $K_d$ of $0.29 \pm 0.02$ mM (Figure 3.9B). This is comparable to the binding affinity for cyanide demonstrated by CYP116B1 (0.54 mM) [298], but significantly tighter than the reported affinity of CYP51B1 (18.8 mM) [57], and demonstrates the variability in binding affinities shown by P450 enzymes for the small cyanide molecule.
Chapter 3: Results I

Figure 3.9: Optical titration of XplA with sodium cyanide. (A) UV/Visible absorption spectra of XplA (5 µM) titrated with cyanide. Aliquots (0.2 µL) of cyanide (1 M stock) were added to a final concentration of 7 mM. The Soret peak in the initial state is at 417 nm and red shifts to 436 nm upon cyanide binding. The inset shows the initial (black) and final (red) spectra from the titration. (B) Binding curve from the titration generated by subtracting the absorption change at 415 nm from the absorption change at 441 nm and plotting these values against the concentration of cyanide added. The data were fitted using Equation 2.1 generating an apparent $K_d$ of 0.29 ± 0.02 mM. The inset shows the difference spectra from which the wavelengths of maximum absorption changes were determined and used to generate the binding curve.
In addition to indole, we also investigated the possibility of XplA binding a series of heterocyclic secondary amines including piperazine, pyrrolidine, and morpholine as substrates. This was due to the discovery of an XplA homologue recently found to be part of a gene cluster in *Gordonia* (sp. KTR9), which also included an XplB-glutamine synthetase (GlnS) fusion enzyme and another P450 that shares 71% amino acid sequence identity with *pipA/morA*-encoded P450s involved in the utilisation of piperazine, pyrrolidine, and morpholine in mycobacteria [299]. Subsequently, binding studies showed all three compounds bound to XplA, albeit as type II (Figure 3.10). In each titration the Soret peak shifts from 417 to 422 nm for morpholine and piperazine binding, and to 423.5 nm for pyrrolidine binding. Furthermore, there is an increase in absorption at the δ peaks at 364 nm for morpholine, and 363 nm for piperazine and pyrrolidine. In addition, diminution of the α peak at 566 nm is also observed in each instance, and is retained only as a shoulder of the β peak, which is located at 540.5, 541 and 545 nm for morpholine, piperazine and pyrrolidine, respectively.
Figure 3.10: Optical titrations of XplA with heterocyclic secondary amines. UV/Visible absorption spectra of XplA titrated with morpholine (A), piperazine (B) and pyrrolidine (C). Similar spectral changes are induced by binding of all three compounds. The Soret peak shifts from 417 nm to 422 nm for morpholine and piperazine binding, and to 423.5 nm for pyrrolidine. There is diminution of the α peak at 566 nm in each titration with an increase in the absorbance at ~363 nm corresponding to the δ peak. The insets show the initial (black) and final (red) spectra from each of the titrations. XplA concentrations used were 6 µM (A), 4 µM (B) and 3.5 µM (C), and ligands were added to final concentrations of 0.5, 10 and 5.5 mM for morpholine, piperazine and pyrrolidine, respectively.

The dissociation constants for morpholine, piperazine and pyrrolidine were determined by fitting the difference in induced absorption changes at 430 and 411 nm against ligand concentration using Equation 2.1 for morpholine, piperazine and pyrrolidine yielding apparent $K_d$ values of $31.8 \pm 0.5$ µM, $970 \pm 40$ µM and $481 \pm 55$ µM for morpholine, piperazine and pyrrolidine, respectively (Figure 3.11). The binding of these heterocyclic amines, and also RDX and imidazole shows that XplA may have specificity for heterocyclic compounds, providing clues to its natural substrate, should it later be determined not to be RDX.
Chapter 3: Results I

Figure 3.11: Determination of the affinity of XplA for heterocyclic secondary amines. The binding curves from titrations of XplA with morpholine (A), piperazine (B) and pyrrolidine (C). Data were fitted using Equation 2.1 for morpholine, piperazine and pyrrolidine, generating \( K_d \) values of 31.8 ± 0.5 µM, 970 ± 40 µM and 481 ± 55 µM, respectively. The insets show the difference spectra, which were used to produce the plots.

The azole family of drugs are common inhibitors of P450s and are used to treat fungal infections in humans [300, 301]. They are five-membered heterocyclic ring molecules containing nitrogen that typically bind tightly to the active site of P450s due to their hydrophobic nature, largely due to substituent groups attached to the azole moiety, with most being imidazole derivatives (as opposed to triazole derivatives as the other main family members) [28, 302]. Optical titrations were performed with a host of azoles including econazole, clotrimazole and ketoconazole amongst others. However, the only compound of this family which induced spectral changes was the smaller imidazole substituted 4-PIM. This suggested that the binding of bulkier, hydrophobic azoles was disfavoured in the constricted environment of the XplA active site. 4-PIM binding induces spectral changes expected of a type II ligand, where the Soret shifts from 417 nm to 423 nm (Figure 3.12A), and concomitant increase in absorption in the...
\( \delta \) peak (362 nm). As observed for the secondary amines described above, there is also a diminution of the \( \alpha \)-peak. A plot of the difference in changes in absorption at 430 and 411 nm produced a binding curve that was fitted using Equation 2.2 (Figure 3.12B) to produce a \( K_d \) value of 0.25 ± 0.04 \( \mu \text{M} \). This is comparable to the binding affinities of 4-PIM demonstrated by most P450s, e.g. BM3 binds 4-PIM with a \( K_d \) value of 0.85 ± 0.45 \( \mu \text{M} \) [275]. The very tight binding observed for 4-PIM is likely to reflect the favourable interactions of the phenyl group of 4-PIM with hydrophobic amino side chain groups in the largely hydrophobic active of XplA, in addition to the Feimidazole ligation. The fact that only 4-PIM binds XplA, and 1-phenylimidazole and 2-phenylimidazole do not is reflective of the constricted nature of the XplA active site, where the orientation of the phenyl group is likely to be dictated by the Fe-imidazole bond, which seemingly introduces steric clashes in 1- and 2-phenylimidazoles.
Figure 3.12: Optical titration of XplA with 4-PIM. (A) UV/Visible absorption spectrum of XplA (5 µM) titrated with 4-PIM. Aliquots (0.1 µL) of 4-PIM (2 mM stock) were added to a final concentration of 8 µM. The Soret peak in the initial state is at 417 nm and red shifts to 423 nm upon 4-PIM binding. The inset shows the initial (black) and final (red) spectra from the titration. (B) Binding curve from the titration generated by subtracting the absorption changes at 411 nm from 430 nm. The data were fitted using Equation 2.2 generating an apparent $K_d$ of $0.25 \pm 0.04$ µM. The inset shows the difference spectrum from which the wavelengths of maximum absorption changes were determined and used to generate the binding curve.

As described earlier, XplA can be purified with imidazole-bound suggesting that binding of this ligand is unusually strong. Imidazole, being a polar molecule, has difficulty accessing the generally hydrophobic active site of P450s and typically binds
Chapter 3: Results I

P450s very weakly in the high mM range, e.g. CYP126 and CYP51B1 from *M. tuberculosis* bind imidazole with apparent $K_d$ values of 2.6 and 11.7 mM, respectively [57, 303]. To determine if XplA bound imidazole unusually tightly for a P450, an optical titration was performed. The spectral changes induced by imidazole binding are very similar with the changes induced by 4-PIM described above. Imidazole binding produces absorption maxima at 361, 424.5 and 543 nm for the $\delta$, Soret and $\beta$ peaks, respectively (Figure 3.13A). The difference in changes in absorption at 430 and 412 nm were used to produce a data plot and fitted using Equation 2.1 (Figure 3.13B) to yield a $K_d$ value of $1.57 \pm 0.04 \ \mu$M. This is extraordinary tight binding of imidazole by a P450 enzyme, when for comparison the *M. tuberculosis* CYP121 binds imidazole with a $K_d$ of $>50$ mM [304]. This explains the difficulty in removing bound imidazole from XplA. The tight binding possibly reflects favourable interactions of imidazole with amino acid residues in the constricted heme active site not observed in other P450s. XplA's affinity for imidazole is stronger than the secondary amines morpholine, piperazine and pyrrolidine, which have similar structures to imidazole and also bind via a nitrogen, is possibly because it is a flatter molecule enabling it to avoid steric clashes which the slightly larger secondary amines cannot, and thus bind the XplA heme weaker compared to imidazole. However, due to the absence of the phenyl group present in 4-PIM, binding of imidazole is not as strong as for 4-PIM.
Figure 3.13: Optical titration of XplA with imidazole. (A) UV/Visible absorption spectrum of XplA (5 µM) titrated with imidazole. Aliquots (0.1 µL) of imidazole (2 mM stock) were added to a final concentration of 12 µM. The Soret peak in the initial state is at 417 nm and red shifts to 424.5 nm upon imidazole binding. The inset shows the initial (black) and final (red) spectra from the titration. (B) Binding curve from the titration generated by subtracting the imidazole-induced absorption changes at 412 nm from those at 430 nm. The data were fitted using Equation 2.1, generating an apparent $K_d$ of 1.57 ± 0.04 µM. The inset shows the difference spectra from which the wavelengths of maximum absorption changes were determined and used to generate the binding curve.
Chapter 3: Results I

3.3.2.4: **Ferrous Imidazole-bound Complex**

During attempts to form CO complexes with purified protein contaminated with imidazole, addition of dithionite produced a peculiar spectrum (Figure 3.14A; inset). The Soret peak is red-shifted to 439 nm, prominent $\alpha$ and $\beta$ peaks also appear at 566 and 538 nm, and a broad feature is visible at $\sim$616 nm. This spectrum is likely from the reduction of imidazole-bound heme to form a ferrous imidazole-bound complex. This is an unusual observation as imidazole typically binds to ferrous heme substantially weaker than to ferric heme, and the binding affinity of imidazole to a P450 with the heme in the ferrous state has not been reported. This is because the additional electron in the ferrous heme iron decreases its propensity to accept the coordinate bond from imidazole, leading to weakened affinity.

An optical titration was performed as previous with imidazole, but this time XplA was firstly fully reduced with dithionite prior to the start of the titration. The titration reproduced the spectrum confirming that imidazole does bind to the ferrous heme of XplA (Figure 3.14A). A plot of the difference in changes in absorption at 440 and 405 nm was used to produce a binding curve and the data fitted using Equation 2.1 (Figure 3.14B) to produce a $K_d$ of $2.73 \pm 0.26$ mM. The $K_d$ of imidazole binding to the ferrous heme of XplA was only detectable owing to the unusually low $K_d$ demonstrated by the ferric heme for imidazole, and thus despite a $10^3$-fold decrease in affinity, the $K_d$ was still within range of measurement using optical titrations. Furthermore, the $K_d$ of imidazole binding to the XplA ferrous heme is lower than the reported $K_d$ values for imidazole binding to the ferric hemes of most P450s, which as described earlier are typically in the mM range.
Figure 3.14: Optical titration of ferrous XplA with imidazole. (A) UV/Visible absorption spectra of XplA (5.5 μM) reduced with dithionite to form ferrous heme, and then titrated with imidazole. Aliquots (0.2 μL) of imidazole (5 M stock) were added to a final concentration of 40 mM. The Soret peak at the start of the titration is at 408 nm and red shifts to 439 nm on imidazole binding. Distinctive peaks at 566 and 538 nm and a broad feature at ~616 nm also appear upon imidazole binding. The inset shows XplA in the ferric (black), ferrous (blue) and ferrous/imidazole-bound (red) states. (B) Binding curve from the titration generated by subtracting the absorption changes at 405 nm from 440 nm, using the data for ligand-free XplA as the baseline value. The data were fitted to Equation 2.1 generating an apparent $K_d$ of 2.73 ± 0.26 mM. The inset shows the difference spectra from which the wavelengths of maximum absorption changes were determined and used to generate the binding curve.
Chapter 3: Results I

3.3.2.5: Sulphur Donor Ligands

During early protein purification experiments in an effort to improve purity, the purification buffer was supplemented with DTT to discourage protein-protein interactions caused by disulphide bridges. Interestingly, XplA was shown to bind to DTT, forming a novel species that likely involves heme coordination by a sulphur ligand with absorption peaks at 453.5, 423.5 and 374 nm (Figure 3.15A). The outer Soret bands result from a hyperporphyrin spectrum due to bis-thiolate coordination of the XplA ferric heme iron, while the 423.5 nm feature results from distal coordination of the iron by a thiol ligand, as described by Sono et al. and by Ullrich et al. [305, 306]. The spectrum is similar to that previously reported for P450cam bound to p-chlorothiophenol, which displayed absorption maxima at 463, 417 and 380 nm, and in slightly different intensities compared to the relative peaks in the spectrum of the XplA-DTT complex. Furthermore, P450cam displayed very high affinity for p-chlorothiophenol, binding with an apparent $K_d$ of 0.65 μM and 3.2 μM in the presence of the substrate camphor [305]. Titration of XplA with DTT and the subsequent plot of the difference in changes in absorption at 457 and 417 nm produced a binding curve that fitted well to Equation 2.1 (Figure 3.17B) to produce a $K_d$ value of 70.5 ± 1.1 μM, which is significantly weaker than the affinity of P450cam for a range of similar sulphur-containing ligands.
Figure 3.15: Optical titration of XplA with DTT. (A) UV/Visible absorption spectra of XplA (4.5 µM) titrated with DTT. Aliquots (0.2 µL) of DTT (100 mM stock) were added to a final concentration of 400 µM. The Soret peak in the initial state is at 417 nm and collapses and shifts to 423.5 nm, with concomitant appearance of a distinct split Soret feature with peaks at 374 and 453.5 nm. The inset shows the initial (black) and final (red) spectra from the titration. (B) Binding curve from the titration generated by subtracting the absorption changes at 417 nm from those at 457 nm, using the starting spectrum as the baseline. The data were fitted using Equation 2.1, generating an apparent $K_d$ of 70.5 ± 1.1 µM. The inset shows the difference spectra from which the wavelengths of maximum absorption changes were determined and used to generate the binding curve.
Subsequent titrations with a range of sulphur-containing compounds also demonstrated binding, with BME producing spectral changes near identical to DTT (Figure 3.16A) with absorption peaks at 453.5, 423.5 and 372 nm, and a merging of the peaks in the Q-band region into a single smooth broad peak with an absorption maximum at ~552 nm and binding with a $K_d$ of $58.5 \pm 3.2 \, \mu$M (Figure 3.16B). XplA will also bind dimethyl sulphide (DMS) (Figure 3.16C) and ethyl methyl sulphide (EMS) exhibiting type II shifts. To determine the dissociation constant for DMS, the difference in absorption changes at 433 and 413 nm versus DMS concentration was fitted using Equation 2.1 to yield a $K_d$ of $81.4 \pm 7.0 \, \mu$M (Figure 3.16D). DMS and EMS likely bind as type II ligands, because in contrast to DTT and BME, the sulphurs cannot exist in the thiolate state in these two compounds and can only ligate to the heme iron in the thiol state. The binding of the generally polar sulphur-containing ligands again show that although the XplA heme active site is largely hydrophobic, binding of small polar ligands is not discouraged, and demonstrative of the general flexibility of P450 enzymes in binding a diverse range of compounds, which is also largely responsible for the keen interest shown for these enzymes.
Figure 3.16: Optical titrations of XplA with sulphur donor ligands. The UV/Visible absorption spectra generated by titration of XplA (~2.8 μM) with BME (A) and DMS (C). Aliquots (0.2 µL) of BME (100 mM stock) were added to a final concentration of 0.5 mM, and DMS (200 mM stock) to 1.4 mM final concentration. The titration with BME induces spectral changes almost identical to those observed with DTT. The split Soret peaks have absorption maxima at 453.5, 423.5 and 372 nm. Spectral changes observed for DMS are consistent with those observed for other type II ligands described earlier, with the Soret shifting to 423 nm, and δ-peak at 365 nm. Binding curves derived from the titrations were fitted to Equation 2.1 generating apparent $K_d$ values of 58.5 ± 3.2 and 81.4 ± 7.0 μM for BME (B) and DMS (D), respectively.

To present further evidence that the split Soret was due to a mixture of the sulphur ligand binding to the heme iron in the thiol and thiolate states, BME ligation to XplA was analysed across a range of pH values from 6.0-8.0. The results in Figure 3.19 show that as the pH increases, the absorption peaks at 453.5 and 372 nm, corresponding to the wavelengths where the bis-thiolate-ligated species absorbs strongest, increase in intensity, whilst the peak at 423.5 nm due to the thiol-ligated species collapses. Conversely, at lower pH values the peak at 423.5 nm is more prominent and the two peaks at 453.5 and 372 nm are diminished. These results are
consistent with the data reported by Sono et al. for the pH-dependent titration of P450cam with 1-propanethiol [305].

![Figure 3.17: pH-dependent binding of BME to XplA.](image)

UV/Visible absorption spectra of ligand-free XplA (5 μM; light blue), and BME-bound XplA at pH values 6.0 (black), 6.5 (red), 7.0 (green) and 8.0 (blue). The 423.5 nm peak is prominent at low pH and decreases in intensity at high pH, with concomitant increase in absorbance of the peaks at 453.5 and 372 nm.

3.3.2.6: Reverse Type I Ligands

Given the ability of XplA to bind and reduce RDX, we investigated whether it would also bind a variety of structurally related compounds, including trinitrobenzene (TNB) and trichlorobenzene (TCB), and induce type I spectral shifts. However, optical titrations showed that none of these compounds induced HS shifts, and minimal spectral changes were observed in the spectra. Interestingly, titration with TNB does show a small increase in the Soret peak, suggesting that TNB may bind as a reverse type I ligand (Figure 3.18). Reverse type I ligands interact with the heme iron differently to type I and II ligands. Moreover, the binding of reverse type I ligands does not displace the 6th water ligand but instead sits on top of it, increasing the population of heme iron in the LS state with concomitant increase in absorbance at 417 nm, as is likely observed in XplA with TNB binding. The δ peak increases due to the TNB, which absorbs strongly in the UV region between 200-300 nm. However, as TNB also shows weak absorbance in the 400-600 nm range (Figure 3.18; inset), these
spectral changes may be due to the TNB directly and not as a result of changes in the heme environment which has been influenced by TNB.

Figure 3.18: Optical titration of XplA with TNB. UV/Visible absorption spectra of XplA (10 µM) titrated with TNB to a final concentration of 500 µM. Minimal spectral changes are induced by the addition of TNB. However, the intensity of the Soret peak at 417 nm increases slightly due to an increase in the LS heme iron species, indicative of the binding of a reverse type I ligand. Above ~300 µM of TNB only very small increases in the Soret peak are observed, and is likely to be due to the TNB. The δ peak continually increases as the concentration of TNB increases, as TNB absorbs strongly in the 200-300 nm UV region. The inset shows the UV/Visible absorption spectra of TNB at 0.5 mM (black) and 1.0 mM (red), with weak absorbance peaks at ~444 and ~563 nm.

3.3.2.7: Determination of the Trinitrobenzene Dissociation Constant

To investigate possible interactions of TNB with XplA and confirm TNB as a reverse type I ligand, a series of competitive binding experiments were performed using TNB and RDX. RDX was titrated to XplA bound to 0.25, 0.5 and 1 mM TNB, respectively (Figure 3.19). The Soret peak shifted from 417 nm to 396 nm in each experiment and $K_d$ values were determined as previously described for RDX titration. In the presence of 0.25 mM TNB, XplA bound RDX with an apparent $K_d$ of $17.1 \pm 0.4 \text{ µM}$. At 0.5 mM TNB, the $K_d$ of RDX binding increased to $23.0 \pm 0.7 \text{ µM}$. Finally, at 1 mM TNB, the $K_d$ of RDX binding to XplA increased to $54.4 \pm 1.9 \text{ µM}$. These observed $K_d$ values of RDX binding were submitted to Equation 2.3 to determine the apparent $K_d$ of TNB binding.
and produced $K_d$ values of 159.1, 240.5 and 194.4 µM to give an average $K_d$ of 198.0 ± 40.8 µM for TNB binding. The reduced affinity of RDX binding to XplA in the presence of TNB demonstrates that TNB enters the heme active site and occupies a position close to the heme iron. RDX binding necessitates dissociation of the bound TNB, hence higher $K_d$ values are observed for increasing concentrations of TNB.

![Figure 3.19: Determination of the affinity of XplA for TNB by competitive optical titrations.](image)

**Figure 3.19: Determination of the affinity of XplA for TNB by competitive optical titrations.** UV/Visible absorption spectra of XplA (~3 µM) bound to TNB at concentrations of 0.25 (A), 0.5 (B) and 1 mM (C), and titrated with RDX (25 mM) to final concentrations of 160, 140 and 240 µM, respectively. Spectral changes induced by RDX binding are observed in all three titrations with the Soret peak shifting from 417 nm to 396 nm. Furthermore, other spectral changes in the Q-band and CT band regions were consistent with those observed for RDX binding in the absence of TNB. Titrations were done as with the earlier titration of XplA with RDX in the absence of TNB, and binding curves (insets) were produced as before from which three observed $K_d$ values were determined (159.1, 240.5 and 194.4 µM). These observed $K_d$'s were submitted to Equation 2.3 to yield an average $K_d$ of 198.0 ± 40.8 µM for TNB binding.

The ligand binding properties of XplA, including binding affinities and absorption maxima are summarised in Table A1 of the addendum.
3.4: Pyridine Hemochromogen Method

3.4.1: Determination of the Heme Extinction Coefficient

Quantification of XplA was initially done by using the extinction coefficient described by Omura and Sato ($\epsilon_{450-490} = 91 \text{ mM}^{-1} \text{ cm}^{-1}$ in the reduced/CO-bound minus reduced XplA difference spectrum) [17]. Due to the variability of the extinction coefficient in different P450 enzymes and incomplete formation of the P450 complex, the pyridine hemochromogen method was used to quantify XplA heme and to determine an extinction coefficient at the oxidised XplA Soret peak by determining the quantity of XplA heme in a known sample of the enzyme. Analysis was done according to the method of Berry and Trumpower [272]. Figure 3.20 shows the pyridine hemochromogen spectrum of XplA and indicates an extinction coefficient of $\epsilon_{417} = 93 \text{ mM}^{-1} \text{ cm}^{-1}$ for the oxidised XplA heme at its Soret peak. Following reconstitution of XplA with FMN, a coefficient of $\epsilon_{417} = 100 \text{ mM}^{-1} \text{ cm}^{-1}$ was used to account for the absorption contribution of the bound flavin. This coefficient was then used to enable estimation of the concentration of the oxidised form of XplA once extensively reconstituted with FMN. The calculated extinction coefficient of the XplA heme is comparable to those calculated for other characterised P450s, e.g. *M. tuberculosis* P450s CYP121 and CYP144 have extinction coefficients of $\epsilon_{416.5} = 110 \text{ mM}^{-1} \text{ cm}^{-1}$ and $\epsilon_{420.5} = 100 \text{ mM}^{-1} \text{ cm}^{-1}$, respectively [57, 307]. The well characterised diflavin reductase-P450 fusion BM3 has an extinction coefficient of $\epsilon_{418} = 95 \text{ mM}^{-1} \text{ cm}^{-1}$ for the full length enzyme, which takes into account the absorption contributions for its FAD and FMN cofactors in their oxidised states [308].
Figure 3.20: Pyridine hemochromogen complex of XplA. UV/Visible absorption spectra of XplA in the ferric state (black), on formation of the pyridine hemichromogen complex (red), and on reduction to the pyridine hemochromogen complex by dithionite (blue). The inset shows a magnification of the 500-600 nm region, highlighting the spectral features of the pyridine hemochromogen complex which absorbs maximally at 556 nm.

3.5: Electron Paramagnetic Resonance Spectroscopy

3.5.1: Identification of the Heme Ligation & Spin State

EPR is a spectroscopic technique used to study chemical species with an unpaired electron. The hemes of P450s in the oxidised state are amenable to this type of spectroscopic analysis. EPR can be used to study the heme environment, including its heme iron ligation and spin states, and is thus a useful tool to analyse the unusual ligand binding properties of XplA. The EPR signals for the ligand-free enzyme (black), and for its complexes with 4-PIM (red) and TNB (green) are shown in Figure 3.21. The native enzyme exhibits a pair of LS species with $g$-values 2.56 ($g_x$), 2.26 ($g_y$) and 1.84 ($g_z$), and 2.49, 2.26 and 1.86. The two species likely arise due to different Cys-Fe-water ligated states, with the high $g_x$ values likely reflecting hydrogen bonding or other interactions of the axial water with nearby residues in the XplA active site, e.g. there is a glutamine (Q438) residue 12 Å away from the heme iron, with its side chain oxygen only 6 Å away from the heme iron [246]. In addition to perturbing the heme environment, at cryogenic temperatures, it may be possible that the side chain of
Q438 is forced (in a proportion of the enzyme) to orient in such a manner that the amine ligation the heme. The $g_z$ values are also high compared to those for other ligand-free P450s, e.g. BM3 has $g$-values at 2.42, 2.26 and 19.2 [309], and P450cam (2.45, 2.26, 1.91) and CYP121 (2.47, 2.25, 1.90) [290, 310]. Binding of 4-PIM leads to a predominant species with $g$-values 2.60, 2.26 and 1.82 and a minor LS species, which is likely due to the retention of water as the 6th ligand and with a near identical set of $g$-values as the native enzyme (2.49, 2.26 and 1.87). The species with $g_z$ 2.60 is likely the 4-PIM-ligated species, and characteristic of an imidazole nitrogen 6th ligand. In the TNB-bound EPR spectrum, only a single LS species with $g$-values 2.49, 2.26 and 1.87 is observed. The $g$-values of this species are also near identical to the LS species observed in the native enzyme. Indeed, it would appear to be the case that TNB reinforces binding of the distal water ligand. Binding of TNB alters environment of the water such that one conformation of the 6th ligand dominates the system. This further demonstrates that TNB binds the active site of XplA as a reverse type I ligand that does not displace the water ligand.

![EPR spectra of native and ligand-bound XplA](image)

**Figure 3.21: EPR spectra of native and ligand-bound XplA.** XplA in the ligand-free state (black) gives an EPR signal showing two LS species with $g$-values 2.56, 2.26 and 1.84, and 2.49, 2.26 and 1.86. The binding of 4-PIM (red), shows a predominant LS species with $g$-values 2.60, 2.26 and 1.82, likely the 4-PIM-ligated species, and retention of a minor water-ligated species with $g$-values near identical to the native enzyme (2.49/2.26/1.82). TNB (green) binding shows a single species with $g$-values 2.49, 2.26 and 1.87, which is near identical to the water-ligated species observed in the native and also 4-PIM-bound enzyme.
More complex EPR spectra are produced for RDX-bound (Figure 3.22A) and DTT-bound (Figure 3.22B) XplA. The RDX-bound spectrum shows a predominant HS species with \( g \)-values 7.65, 3.95 and 1.77 with at least four other LS species, three of which have \( g \)-values of 2.54/2.26/1.85, 2.48/2.26/1.86 and 2.41/2.26/1.89. The four LS species are likely to be different hexa-coordinated water-ligated forms of XplA, where interactions of the RDX with the water and/or RDX-dependent perturbations to its environment and ligation geometry give rise to the different species. A further \( g_x \) feature is seen at 1.93. Its corresponding \( g_z \) is less defined, but according to the Bohan formula should be located at 2.33 (see asterisk in Figure 3.22A) [311]. This final component of the spectrum (2.33/2.26/1.93) is also assigned to a discrete form of water-ligated XplA interacting with RDX in the active site. The DTT-bound spectrum shows two species with the same sets of \( g \)-values as the ligand-free XplA, and indicates forms in which the water ligand is retained. In addition, a third ferric heme EPR signal with broad \( g_x \) feature (~2.38-2.42) with \( g_x \) at 1.91 arises due to DTT binding as axial ligand in both its thiol (higher end \( g_z \)) and thiolate (low end \( g_z \)) forms. The \( g \)-values of the thiol and thiolate ligated species for DTT-bound XplA are very similar to the \( g \)-values reported for P450cam bound to DMS and hydrogen sulphide, which have \( g_z \) features at 2.42 and 2.39, respectively [305]. The radical species with \( g \approx 2 \) is the FMN SQ, which arises due to a proportion of the FMN being reduced by DTT to the SQ state.
Figure 3.22: EPR spectra of RDX- and DTT-bound XplA. (A) RDX-bound XplA gives rise to an EPR signal with multiple species, including a major HS species with $g$-values 7.65, 3.95 and 1.77. Four other LS signals are also detected, that likely arise due to retention of the axial 6th water ligand and its subtle changes in environments/positions as a consequence of binding of RDX. (B) A trio of EPR signals are detected in the DTT-bound form of XplA, two of which are identical to the LS species observed in the native enzyme. The third species with broad $g_z$ feature (~2.38-2.42) and $g_x$ at 1.91 emerges due to DTT binding in both the thiol and thiolate states.
Chapter 3: Results I

3.6: Magnetic Circular Dichroism Spectroscopy

To further characterise the heme environment of XplA, substrate-free and ligand-bound XplA was analysed by MCD spectroscopy. Along with EPR, MCD studies of ferric P450s are informative in relation to the heme iron spin state and ligation states [312]. Unique ligands produce unique MCD spectra that can be used as a fingerprint for that ligand type.

3.6.1: MCD Analysis of XplA

The MCD spectra of XplA in NUV/Visible and NIR for the ligand-free enzyme (black), and for its complexes with imidazole (red), BME (green) and RDX (blue) are shown in Figure 3.23. The native enzyme displays spectral features consistent with the expected Cys/H$_2$O axial ligation of the ferric heme. The pattern of bands between 300 and 700 nm indicates the existence of solely LS conformers [313], whilst the position of the CT$_{LS}$ band (1190 nm) in conjunction with the reduced peak to trough intensity at 418 nm defines thiolate as one of the ligands to the heme [314]. The red shift of the Soret band to 425 nm in the imidazole complex is indicative of complete displacement of water by a nitrogenous ligand, the 425 nm peak to trough intensity and position of the CT$_{LS}$ band being consistent with thiolate remaining as the proximal ligand. Binding of BME gives the expected hyperporphyrin spectrum of the ferric complex with a negative Gaussian feature centred on 374 nm and a derivative at 454 nm arising from bis-thiolate coordination, with the 425 nm intensity arising from thiolate/thiol coordination [305]. Dissociation of water to give HS penta-coordinate heme on binding of RDX is evidenced by the appearance of negative Gaussian intensity at 650 nm. However, persistence of the derivative band at 419 nm indicates the presence of both high- and low- spin conformers presumably due to incomplete dissociation of the distal water in the sample.
Figure 3.23: MCD studies of XplA. MCD spectra in the NUV/Visible (A) and NIR (B) regions are shown for XplA. Spectra were collected using 250 μM XplA, with ligands imidazole, BME and RDX in concentrations of 0.5, 5 and 0.5 mM, respectively. The figures show XplA in ligand-free (black), imidazole-bound (red), BME-bound (green) and RDX-bound (blue) forms. The NUV/Visible MCD spectrum of native enzyme is consistent with a LS Cys thiolate/H$_2$O-coordinated P450, whereas RDX binding leads to development of a HS population as evident from development of the 650 nm band. Imidazole and BME binding give signals consistent with formation of LS distal N- and S-coordinated XplA (respectively). A NIR CT band assigned to a porphyrin to ferric CT transition of the XplA heme is located at approximately 1190 nm. CT bands shifts consistent with Cys-BME (green), Cys-imidazole (red) and Cys-RDX (blue) axial ligands are observed. With the peak ~900 nm assigned to the CT band of HS ferric heme.
In the NIR region, LS ferric hemes give a porphyrin to ferric CT transition, appearing as a positively signed MCD band [313]. This can be diagnostic in identifying axial ligands to a heme iron. Previous studies of other bacterial P450s have assigned this band at ~1125 nm for CYP121 [290], 1065 nm for BM3 [315] and at 1110 nm for the CYP51B1 [57, 290, 315]. In the case of XplA, the NIR MCD band is at ~1190 nm (Figure 3.23B; black). This is a slightly longer wavelength than for the other P450s previously studied, but the NUV/Visible MCD and UV/Visible absorption spectra of the same protein sample are consistent with cysteine thiolate as the proximal ligand to the heme iron, and with the XplA being imidazole-free. The NIR MCD data suggest that the heme iron t₂g orbital is closer in energy to the porphyrin orbitals than is normally the case for Cys thiolate coordination, possibly as a consequence of alterations in the environment of the proximal ligand. Such changes would be consistent with a less “electron donating” Cys thiolate and possibly leading to a relatively positive P450 heme iron reduction potential of the substrate-free XplA Fe³⁺/Fe²⁺ redox couple. Binding of imidazole shifts the CTₐₕ band to ~1215 nm (Figure 3.23B; red), similar to the imidazole-bound complex of BM3 (1180 nm) and fluconazole-bound complex of CYP51B1 (1165 nm), which also experience CTₐₕ shifts to a longer wavelength from 1065 and 1110 nm for BM3 and CYP51B1, respectively [57, 315]. Due to the longer wavelength of the ligand-free enzyme, the imidazole-bound CTₐₕ band of XplA is also slightly longer than for BM3 and CYP51B1. The CTₐₕ band shifts to a shorter wavelength when complexed to BME and is centred at 1120 nm (Figure 3.23B; green), which is likely to be diagnostic of a P450 bound to sulphur donor ligands.

3.7: Redox Potentiometry

3.7.1: Determination of the XplA Heme Iron Reduction Potential

XplA contains two redox centres with three redox couples, the Fe³⁺/Fe²⁺ of the heme and the OX/SQ and SQ/HQ couples of the FMN moiety. Determination of the redox potentials of XplA indicates how readily each redox active species accepts electrons and subsequently be reduced. To determine the midpoint reduction potentials of these three redox couples, redox titrations were performed under anaerobic conditions using the methods of Dutton as described in Materials & Methods (2.2.17)
The redox potentials of XplA were determined by progressively decreasing the redox potential of the solvent by titrating with dithionite. The changes in redox potential of the surrounding solvent induced spectral changes of XplA as a consequence of heme iron reduction, with a decrease in the intensity of the Soret peak for the oxidised ferric heme iron at 417 nm, and an increase at 408 nm, corresponding to the reduced, LS ferrous heme iron, and strongly suggestive of the retention of a thiolate ligand to the iron (Figure 3.24A). Spectral changes at ~460 nm and ~630 nm are also observed, these changes are induced by flavin reduction and subject to further discussion in Chapter 5 (5.6.5).

The absorption change at 417 nm was plotted against the applied potential and the data fitted to a 1-electron Nernst equation to determine the midpoint reduction potential for XplA heme iron in the absence of substrate (Figure 3.24B). The heme reduction potential was determined to be -268 ± 5 mV. This reduction potential is high compared to that for e.g. the M. tuberculosis cyclodipeptide oxidase CYP121, which has a reduction potential for substrate-free enzyme of -467 mV [304], and CYP144 also from M. tuberculosis, which has a reduction potential of -355 mV [307]. The high potential of the XplA heme explains the tendency of XplA to be readily reduced in an aerobic environment by dithionite.

The reduction potential of substrate-bound XplA could not be determined, as dithionite directly reduces the substrate RDX. However, potentiometric studies of some well characterised P450s demonstrate that, on substrate binding, the redox potential of the heme typically increases by >100 mV [57, 274]. Therefore, it can be expected that the heme redox potential of RDX-bound XplA will likely increase to ~150 mV. This is significantly higher than the redox potential of arachidonic acid-bound BM3 (-289 mV) [14], but comparable to that observed for CYP142 which displays a potential of -192 mV with cholest-4-en-3-one bound, a remarkable increase of ~220 mV from the substrate-free enzyme [294]. This high potential of the XplA heme suggest it has evolved to rapidly accept electrons from redox partners to reduce the RDX substrate, and possibly points to relatively high potentials for the FMN moiety aswell.
Figure 3.24: Spectroelectrochemical redox titration of ligand-free XplA.  (A) Spectral changes of XplA (8 µM) induced by reduction with dithionite. Arrows indicate directions of absorbance change during the reductive phase of the titration. At ~470 nm (near the absorption maximum for oxidised FMN) there is a decrease in absorbance in the early phase, reflecting conversion of FMN from oxidised through SQ to HQ. At more negative potentials, there is an increase in absorption in this region due to heme iron reduction. In the region at ~550 nm there is an initial increase in absorption due to FMN SQ formation, followed by a decrease as the SQ is reduced to FMN HQ, and then a final increase as heme is reduced and the Q-band absorption becomes more prominent. Spectral changes at 417 nm and 408 nm are consistent with a ferric to ferrous heme transition. The inset shows the initial (black line) and final spectra from the titration. (B) Plot of $A_{417}$ versus potential fitted using the Nernst equation to give a heme iron Fe$^{3+}$/Fe$^{2+}$ midpoint potential (versus NHE) of $E = -268 \pm 5$ mV.
3.8: Stopped-flow Spectroscopy

A stopped-flow instrument is an apparatus that allows the rapid mixing of two or more different solutions (such as an enzyme and its substrate) in order to study their reaction kinetics on a timescale of a few milliseconds and upwards. The reaction is usually observed by means of a form of spectroscopy, typically UV/Visible or fluorescence spectroscopy. Stopped-flow spectroscopy is used to study reaction kinetics too rapid to be studied using steady-state methods, such as the binding of diatomic gases like CO and O₂ to the ferrous heme of P450s.

3.8.1: Determination of the XplA CO Dissociation Constant

CO is referred to as an essentially irreversible inhibitor of many hemoproteins. It is therefore surprising to observe that a P450 substrate, which often binds more weakly than inhibitors, was able to displace CO from the carbon-monoxy complex. To investigate CO binding, the kinetics of Fe(II)CO complex formation (at 446 nm) was studied using stopped-flow mixing of dithionite-reduced XplA heme domain with CO at various concentrations and described in Materials & Methods (2.2.18). Figure 3.25 shows a typical reaction kinetic transient reporting on the increase in heme absorption at 446 nm as the Fe(II)CO complex forms. ΔA₄₄₆ reaction transients were fitted accurately using a monoexponential function to produce a set of k_obs values. The inset shows a plot of the rate constants for CO binding to XplA ferrous heme iron against the concentration of CO, and this shows a linear dependence of the rate of CO-complex formation on the concentration of CO. From the data, the rates of CO association (k_on) and dissociation (k_off) were determined from the gradient of a linear fit of the observed rate constants for CO binding (k_obs) versus the applied CO concentration, and from the y-axis intercept of the fit line, respectively. Thereafter, the K_d was calculated as 2.64 ± 0.51 μM by dividing k_off (19.1 ± 3.6 s⁻¹) by k_on (7.24 ± 0.12 μM⁻¹ s⁻¹). This is comparable to the binding affinity shown for RDX by XplA in the ferric state, and shows that ferrous XplA binds CO only moderately tighter than it does its substrate RDX. Furthermore, XplA shows weakened affinity for CO compared to other characterised P450s, e.g. adrenal P450ssc binds CO with an apparent K_d of 0.3 μM at 25 °C [316], and is likely due the high dissociation rate of CO demonstrated...
Chapter 3: Results I

by XplA. Presumably this relatively high $K_d$ for CO binding for a P450 contributes to the observation that RDX is able to oxidise the Fe(II)CO complex.

![Graph of CO binding kinetics for XplA Fe(II)CO complex formation](image)

**Figure 3.25: Kinetics of the formation of the XplA Fe(II)CO complex.** The Fe(II)CO complex of XplA was formed by anaerobic stopped-flow mixing of ferrous XplA with various concentrations of dissolved CO. Absorption change was monitored at 446 nm, the absorption maximum for the Fe(II)CO complex. The main panel shows a typical monophasic $\Delta A_{446}$ reaction transient (recorded at 11.6 μM CO, $k_{obs} = 98.5 \pm 0.2$ s$^{-1}$). The inset shows a plot of the observed rate constants ($k_{obs}$ values) for Fe(II)CO complex formation against the relevant CO concentration, fitted using a linear function. Data fitting provided parameters of $k_{on} = 7.24 \pm 0.12$ μM$^{-1}$ s$^{-1}$ for the 2$^{nd}$ order rate constant for CO association, $k_{off} = 19.1 \pm 3.6$ s$^{-1}$, and a $K_d$ value of 2.64 ± 0.51 μM for CO binding derived from $k_{off}/k_{on}$.

3.8.2: Oxyferrous Complex Characterisation

Researchers have been drawn to P450s owing to their ability to perform oxidative chemistry. However, it appears that XplA has evolved to perform reductive chemistry under the evolutionary pressure of RDX in the environment [257, 283], and that the proposed mechanism of RDX degradation proceeds without oxidative chemistry [249, 258]. We sought to establish whether XplA was still able to bind dioxygen and perform traditional oxidative chemistry associated with P450s. Previous studies of XplA showed differences in product formation from RDX under anaerobic and aerobic conditions, with nitrite and formaldehyde accompanied by predominantly MEDINA anaerobically, or by NDAB in aerobic reactions [106]. While it was considered
unlikely that XplA would directly hydroxylate RDX [106, 309], we examined the ability of the P450 to form an oxyferrous complex using a stopped-flow approach.

The observation of indigo formation in the cell cultures, presumably because they are expressing xplA suggest that it is capable of catalysing hydroxylation reactions. In addition, XplA has a relatively high redox potential which may help stabilise the oxyferrous complex, as electron transfer to the oxygen is disfavoured by the high potential of the heme iron by comparison with other P450s of more negative heme iron potential [317]. Furthermore, as we had the intact XplA and XplA-HD constructs, experiments using the intact XplA were also used, since these may also allow us to spectrally visualise the ferric-peroxo or protonated ferric hydroperoxide species formed on second electron delivery to the oxyferrous complex from reduced FMN in the XplA flavodoxin. The XplA oxyferrous complex was characterised by mixing oxygenated buffer (100 μM O₂) in one syringe with deoxygenated ferrous XplA heme domain in the other syringe in a glove box environment at 4 °C. Entire spectra were accumulated using a photodiode array attachment, and demonstrated clearly the formation of an oxyferrous complex of XplA with the Soret band at 424.5 nm and a strong Q-band feature centred at 557 nm (Figure 3.26). The UV/Visible absorption spectrum of the XplA oxyferrous complex is similar to that previously reported for the heme domain of BM3, which displayed absorption maxima at 422 and 559 nm, respectively [318]. The experiment was also repeated for the full length enzyme, however no differences in the spectrum of the oxyferrous complex were observed.
Figure 3.26: Characterisation of the XplA oxyferrous complex. UV/Visible absorption spectra of XplA-HD (5 μM) in the Fe(II)O2 (red) and ferric states (black). The oxyferrous complex was observed by mixing ferrous XplA-HD (10 μM) with oxygenated buffer (100 μM). The XplA-HD oxyferrous complex has a Soret peak at 424.5 nm, and a distinctive Q-band feature at 557 nm.

3.8.3: Kinetics of Oxyferrous Complex Formation & Collapse

Single wavelength stopped-flow absorption spectroscopy was then used to determine rate constants for O2 binding to ferrous XplA at different concentrations of the gas (kobs). Reaction transients were monophasic and Figure 3.27A shows the linear dependence of kobs on [O2], which was fitted to give a 2nd order rate constant for O2 binding (kon) of 6.64 ± 0.15 s⁻¹ μM⁻¹, and an apparent rate of O2 dissociation (koff) of 2.25 ± 1.97 s⁻¹, resulting in an O2 dissociation constant of 0.34 ± 0.30 μM at 4 °C. This is comparable to the Kd previously reported for the microsomal P450LM (0.68 μM), determined using the same methods by Rösen and Stier [319]. Although the CO and O2 association rate constants for ferrous XplA are comparable, the increased dissociation rate constant of CO when compared to O2 is responsible for the weaker binding observed for CO in XplA.
Figure 3.27: Kinetics of XplA oxyferrous complex formation and collapse. The Fe(II)O₂ complex of XplA was formed using identical methods to the Fe(II)CO complex. Absorption change was monitored at 435 nm, near the absorption maximum for the Fe(II)O₂ complex. Panel (A) shows a plot of the observed rate constants ($k_{\text{obs}}$ values) for Fe(II)O₂ complex formation against the relevant O₂ concentration, fitted using a linear function. Data fitting provided parameters of $k_{\text{on}} = 6.64 \pm 0.15 \text{ s}^{-1} \text{ μM}^{-1}$ for the 2nd order rate constant for CO association, $k_{\text{off}} = 2.25 \pm 1.97 \text{ s}^{-1}$, and a $K_d$ value of $0.34 \pm 0.30 \text{ μM}$ for O₂ binding derived from $k_{\text{off}}/k_{\text{on}}$. Panel (B) shows a plot of the oxyferrous complex collapse by plotting the $A_{435}$ against time. The data were fitted to a single exponential to yield an apparent decay rate of $1.98 \pm 0.02 \text{ s}^{-1}$. 
Chapter 3: Results I

The stability of this complex was analysed by determining its rate of decay, by plotting the absorption at 435 nm, where the difference in absorption between the ferrous and oxyferrous complex is largest against time. Fitting of the data a single exponential function produce an apparent decay rate constant of $1.98 \pm 0.02 \text{s}^{-1}$ at 4 °C (Figure 3.27B). This rate is significantly higher than that reported for P450cam (0.01 s$^{-1}$ at 25 °C) [320], and also to P450 aromatase (0.21 s$^{-1}$ at 25 °C and 0.7 s$^{-1}$ at 37 °C) [321], and indicate that the stability of the oxyferrous complex does not correlate well with the reduction potential of the heme, as the Fe$^{3+}$/Fe$^{2+}$ midpoint reduction potential of P450cam is -330 mV [322], which is lower than that of XplA Fe$^{3+}$/Fe$^{2+}$. The stability of this complex is therefore less stable than those previously described, but sufficiently stable to be characterised using stopped-flow techniques, e.g. although the Fe(II)CO complex of P450scc was successfully characterised, the binding of O$_2$ to substrate-depleted enzyme was too rapid to measure at 7 °C [316]. Due to the transient nature of the oxyferrous complex of the P450s, the kinetics of the formation and decay of this and other intermediates of the P450 catalytic cycle have not been extensively studied.

3.9: Summary

The novel cytochrome P450-flavodoxin enzyme XplA was expressed and purified to homogeneity. Subsequent characterisation of XplA using UV/Visible, EPR and MCD spectroscopy revealed unusual spectroscopic and ligand binding properties. Binding of the RDX substrate was observed using UV/Visible spectroscopy and confirmed by EPR and MCD spectroscopy. However, the binding affinity was ~10-fold tighter compared to the previously reported affinity [106]. Furthermore, the spectral transitions induced by RDX binding in our titrations were consistent with those expected for substrate binding to a P450, which contrasted with the non-standard spectral changes described by Jackson et al. for RDX binding [106]. It is reasonable to assume that the larger $K_d$ reported previously was a result of the retention of imidazole in samples used in titration experiments with RDX, given the unusually strong affinity of XplA for imidazole, which was shown to bind XplA with an apparent $K_d$ of 1.57 μM. This is at least 10$^3$-fold tighter than the P450s which had previously showed the strongest binding to imidazole, which only demonstrated $K_d$'s in the low mM range. Due to the polar nature of imidazole, binding of this molecule is typically
very weak in most P450s. In addition, the affinity of XplA for imidazole is so strong that binding to the reduced enzyme was also observed, albeit binding was a lot weaker than for the enzyme in the ferric state. The UV/Visible absorption spectrum of the imidazole-bound ferrous XplA was unique and has not previously been reported with absorption peaks corresponding to the heme at 439 nm and other absorption maxima at 538 and 566 nm.

Other peculiarities observed for this enzyme include binding to a series of sulphur donor ligands, including DTT and BME, to produce spectra mirroring those reported for P450cam almost 30 years earlier [305]. In addition, the common buffer constituent, Tris, was also shown to bind to XplA, exacerbating the problem of purifying XplA in a LS form with water as the 6th ligand, as Tris-based buffers were commonly used. Binding of several common P450 inhibitors (the azole family of drugs) was not observed, likely as a consequence of the constricted environment of the XplA active site. Interestingly, it was also shown that RDX was able to displace CO from the Fe(II)CO complex. The bound CO is likely displaced as a result of the oxidation of the ferrous heme by RDX following CO dissociation. CO cannot then re-bind to the ferric heme leading to the rapid collapse of the P450 complex. No substrate-like binding was observed for any other compound tested, including the structurally related explosive molecules TNB and TCB. However, reverse type I binding was observed for TNB. Furthermore, binding of the tryptophan breakdown product, indole, was not demonstrated. However, indole hydroxylation is observed in cells expressing xplA, and this apparently demonstrates the diverse substrate selectivity of XplA, and indicates that it is still capable of catalysing traditional P450 reactions involving oxygen activation and insertion into substrates.

MCD and EPR data collected for this enzyme supported the UV/Visible data and confirmed the unusual spectroscopic and ligand binding properties of XplA. The EPR spectrum of native XplA showed the presence of two LS species with higher $g_z$ (2.56 and 2.49) features than those commonly observed for other native P450s. Most P450s only give one EPR signal due to the water-ligated species. The two LS species likely arise due to interactions of the 6th water ligand with amino acids in the active site. Moreover, the LS species with higher $g_z$ (2.56) feature is displaced in favour of the water-ligated species upon TNB binding, perhaps indicating that it is itself a
Chapter 3: Results I

water-ligated species and that TNB has the effect of organising the distal water to occupy a single orientation on the heme iron.

The thermodynamic properties of the native enzyme were also characterised following reconstitution with FMN, and the potential of XplA heme iron was determined to be -268 mV, significantly more positive than the values for BM3 and CYP121 (~-427 and -467 mV, respectively) [14, 304]. This potential is expected to increase upon RDX binding by at least 100 mV. However, RDX was susceptible to reduction by dithionite and the reduction potential of RDX-bound XplA could not be measured. The relatively positive redox potential of XplA allows it to rapidly reduce the RDX molecule by efficiently transferring electrons through the redox partners to the heme iron and onto the RDX. The extinction coefficient of the oxidised heme of XplA was also determined using the pyridine hemochromogen method, and an extinction coefficient of $\varepsilon_{417} = 93 \text{ mM}^{-1} \text{ cm}^{-1}$ for heme Soret peak was determined. This was adjusted to 100 mM$^{-1}$ cm$^{-1}$ to account for the absorption contributions of the flavin, when fully incorporated.

Although it has been proposed that XplA evolved to perform a purely reductive role, its ability to bind oxygen was demonstrated using stopped-flow methods. Furthermore, the observation of indigo formation in cell cultures expressing XplA gave further indication that the enzyme could at least catalyse hydroxylation reactions, as the hydroxylation of indole to 2-hydroxyindole is a requisite of indigo production. Previously, XplA was demonstrated to catalyse the reductive degradation of RDX in the absence of oxygen [106], and also aerobically, albeit the rate was slower in the presence of oxygen. The confirmation of oxygen binding provides an explanation for the discrepancy in rates observed in aerobic and anaerobic environments. The slower rate observed aerobically is due to non-productive binding of oxygen to the ferrous heme, which interferes with electron transfer to RDX, although the binding of RDX is unlikely to be perturbed by oxygen. The UV/Visible absorption spectrum of the oxyferrous complex of XplA, and its binding affinity for oxygen is similar to those reported for other P450s, with the Soret peak red shifted to 424.5 nm and a single sharp peak appearing in the Q-band region at 557 nm. The $K_d$ was determined to be 0.33 µM. In addition, the kinetics of oxygen binding was not been significantly altered, despite the absence of the conserved acid/alcohol amino acid pair which is key to the functionality of most P450s. Although, the absence of
these residues, which are involved in the proton relay network, is more likely to affect the subsequent intermediates following the oxyferrous complex in the P450 catalytic cycle. Attempts to characterise the twice reduced ferric-peroxo and protonated ferric hydroperoxide intermediates were unsuccessful using the intact enzyme, and only the same oxyferrous complex was observed as for the heme domain. These results show that although the primary role of XplA may be to function as a reductase, it is still capable of oxygen chemistry like all P450s and raises the possibility that XplA has a yet undefined role with a natural substrate other than RDX.

The same stopped-flow techniques were also used to determine the $K_d$ for CO binding, and although the 2nd order rate constant for CO binding was greater than that for O$_2$ binding, the CO dissociation rate constant was significantly faster than that for O$_2$, resulting in a $K_d$ of 2.64 µM for CO. This is relatively weak compared to previously reported binding affinities for CO by other P450s, which maybe a contributory factor in the observation of rapid RDX-dependent oxidation of CO-bound ferrous heme.

In this chapter, non-standard spectroscopic and ligand binding properties of XplA were elucidated, revealing novel features of a P450 enzyme. Furthermore, the protein was isolated in a pure and stable form using a two step chromatographic procedure. This facilitates structural characterisation of XplA, which is the subject of analysis in Chapter 4. In addition, during the purification process, unusual properties of the FMN-binding domain were also observed, and this is subject to analysis in Chapter 5.
Chapter 4: Structural Characterisation of XplA
Chapter 4: Results II

4.1: Introduction

The crystallisation of multi-domain proteins has traditionally been a difficult process, which often ends in the failure to obtain crystals and is a major factor in the rarity of reported structures of such proteins determined using X-ray crystallography. Therefore, several obstacles stand in the pursuit of obtaining crystals of the full length XplA. Furthermore, the structural characterisation of a P450-redox partner fusion enzyme has never been reported. However, XplA represents a simpler P450-redox partner fusion enzyme to those previously described with only two domains, the heme- and FMN-binding domains, which may be more amenable to crystallogenesis. The only other described native P450 fusion protein comparable to XplA, is MCCYP51FX, which consist of the sterol 14α-demethylase CYP51 fused to a small ferredoxin protein [225]. Structural studies of this fusion enzyme have not been described. An engineered construct of BM3 containing the heme- and FMN-binding domains was successfully crystallised but, as described in Chapter 1, the structure obtained from these crystals showed the linker region between the two domains had proteolysed [211].

The structural characterisation of fusion enzymes is appealing, as these structures can reveal detailed information which the structures of the isolated enzymes cannot provide, e.g. they may enable the identification of important residues at the interacting interface, or allow the proposal of electron transfer pathways. For example, the structural characterisation of CPR revealed detailed information on the FAD and FMN environments, e.g. the structure showed the two redox cofactors were only 4 Å apart, and provided insight into their interactions. Furthermore, the recent determination of the crystal structure of covalently cross-linked PdR-Pdx, and previously of an AdR-Adx covalently cross-linked structure, supports previously determined experimental data that showed interactions were driven primarily by electrostatic and hydrophobic interactions [231, 232]. However, these structures were only determined because the proteins were locked in a rigid conformation by the covalent linkages, and may not be representative of the native proteins.

Given the low probability of obtaining crystals of the full length protein, the P450 and flavodoxin domains of XplA were separated and expressed individually. The isolation of the P450 domain would improve the chances of obtaining a structure of RDX-bound XplA. The elucidation of the structure of XplA bound to RDX, or a breakdown
product thereof, may provide mechanistic details of RDX breakdown to support the proposed degradation pathways [106, 249, 323]. Furthermore, a structure of the XplA-HD has been reported, demonstrating the crystallisability of XplA-HD [246]. The isolation of the XplA flavodoxin would also facilitate structural and other biochemical characterisation of this flavodoxin protein, which otherwise may have been difficult when fused to the P450. This chapter reports on the X-ray crystallographic studies of the full length XplA and the isolated heme domain, and on pre-crystallisation studies of both enzymes to determine their suitability for crystallisation.

4.2: Molecular Biology of the XplA-P450 Domain

To generate a plasmid construct expressing the XplA P450 domain (XplA-HD, residues 159-552 of XplA), an Ndel restriction site was engineered in the linker region separating the FMN and heme domains. To achieve this, xplA-HD was generated by first using oligonucleotide primers XplA-HDF (5’-GTCTGTTGTTGGTCACCATGACTGCACCTC-3’) and XplA-HDR (5’-GAGGCCGCAGTCATAATGGACTTGACCACGAC-3’), where non-complementary nucleotides introduced are in bold and the new site generated for the restriction enzyme Ndel is underlined. The mutagenesis PCR experiment was done using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene), following the manufacturer's protocol. The FMN domain was excised from plasmids carrying the mutation by digesting with Ndel and the vector re-ligated using the complementary ends produced by the Ndel digest to produce the construct pET15(xplA-HD). Excision of the FMN domain from the mutagenised xplA gene was confirmed by analysing the digest on a DNA gel where a DNA fragment of the approximate size of xplA-FMN (474 bp) was observed (Figure 4.1). Successful mutation and ligation were confirmed by sequencing plasmids purified from successful transformants, and by restriction digestion of the plasmids using Ndel and BamHI, and visualisation of a DNA fragment in a DNA gel of a size expected for xplA-HD (1,185 bp) (Figure 4.2).
Figure 4.1: *NdeI* restriction digestion of mutagenised pET15b(xplA). (A) 1.4% (w/v) agarose DNA gel showing a 15 µL *NdeI* restriction digestion of pET15b(xplA) containing an introduced second *NdeI* restriction site following PCR mutagenesis. The lanes from left to right show; DNA ladder (1), undigested plasmid (2) and *NdeI* digested plasmid (3). The digest results in two fragments, the linearised vector of 6,893 bp in size and the small insert excised of 474 bp. (B) 1.4% agarose DNA gel showing the same digest as gel A in a 40 µL reaction volume. The *NdeI* digested vector, free of the *xplA-FMN* domain, was excised and re-ligated to form the pET15b(xplA-HD) construct. The lanes from left to right show DNA ladder (4), undigested plasmid (5) and *NdeI* digested plasmid (6).

Figure 4.2: *NdeI* and *BamHI* double restriction digestion of pET15b(xplA-HD). 0.8% (w/v) agarose DNA gel showing *NdeI* and *BamHI* restriction digestions of plasmids recovered from successful transformants of re-ligated pET15b(xplA-HD). The lanes from left to right show; DNA ladder (1), undigested plasmid 1 (2), *NdeI* and *BamHI* digested plasmid 1 (3), undigested plasmid 2 (4), *NdeI* and *BamHI* digested plasmid 2 (5), undigested plasmid 3 (6) and *NdeI* and *BamHI* digested plasmid 3 (7). The digests confirm the successful re-ligation of *NdeI* treated PCR mutagenesis products and shows the separated *xplA-HD* gene insert running at its expected size 1,185 bp.
4.3: Expression & Purification of XplA-HD

The pET15b(xplA-HD) construct was transformed into HMS174 (DE3) cells and heterologous recombinant protein expression was induced as with the full length protein. The construct was also transformed into other E. coli strains, and similar levels of XplA-HD expression were observed in all strains. XplA-HD was purified using Ni-NTA affinity chromatography as with the full length enzyme, at which stage the protein was pure (Figure 4.3). For crystallogenesis experiments, XplA-HD was purified further using size-exclusion chromatography as a final polishing step and to buffer exchange into 50 mM HEPES, 100 mM NaCl (pH 7.0). Approximately 24 mg of homogenous protein was attained from 12 L of culture (~2 mg/L).

Purified XplA-HD displays a UV/Visible absorption spectrum characteristic of most P450s and is compared to the spectrum of the intact XplA in Figure 4.4. Highly purified XplA-HD achieves an $A_{418}/A_{280}$ ratio of 1.47:1, and has a smaller 280 nm peak compared to the intact enzyme due to its smaller size. In comparison, the purest XplA fractions achieve a ratio no greater than 1:1. As with the full length protein, extensive dialysis of XplA-HD was also required to remove the bound imidazole. The Soret peak of the dialysed protein is at 417 nm, and the $\alpha$ and $\beta$ peaks at 566 nm and 540 nm,
respectively. These spectral features are identical to those of the full length protein, and are not affected by the absence of the FMN cofactor. However, there is substantial bleaching at ~360 nm and between 450-500 nm compared to the full length protein corresponding to regions where absorbance of the oxidised FMN is strongest.

![Spectroscopic characterisation of purified XplA-HD](image)

**Figure 4.4: Spectroscopic characterisation of purified XplA-HD.** UV/Visible absorption spectrum of XplA-HD (8 µM), following Ni-NTA affinity and size-exclusion chromatography purification (black). Shown for comparison is the spectrum of full length XplA (red), which has a larger 280 nm protein peak due to its larger size, and increased absorbance contributions of the flavin between 450-500 nm and ~360 nm, respectively. The wavelengths of spectral features corresponding to the heme Soret, α and β peaks are identical to those of XplA.

### 4.4: Biophysical Characterisation of XplA & XplA-HD

The P450 domain of XplA was successfully isolated, expressed and purified. Removal of the flavodoxin domain from XplA was not detrimental to the expression and, more importantly, solubility of XplA-HD, and sufficiently large quantities of soluble protein was recovered for further biophysical and structural studies. To determine the suitability of XplA and XplA-HD for crystallographic studies, the biophysical properties of the proteins were analysed using a variety of biophysical techniques to determine the protein's stability, folding and polydispersity index. Therefore XplA
and XplA-HD, were analysed using the Thermofluor assay, CD spectroscopy and MALLS.

4.4.1: *Thermofluor Assay*

To aid crystallographic studies of a protein, it is often advantageous to identify buffer conditions which minimise protein aggregation and improve protein stability. Recent structural characterisation of proteins has incorporated a pre-crystallographic tool to complement MALLS called the Thermofluor assay to identify buffer conditions which maximise protein stability [324, 325]. The assay uses a PCR cycler to perform high-throughput screening of buffer conditions in 96-well plates, which makes it especially useful in identifying suitable buffers from commercially available crystallisation screens. The assay utilises the property of a fluorescent dye called SYPRO Orange, whose fluorescence increases when exposed to a hydrophobic environment [326]. The thermally induced unfolding of a protein exposes the normally buried hydrophobic core of a folded protein, thereby increasing the hydrophobic environment. Therefore, the thermal unfolding of a protein can be followed by monitoring the fluorescence of SYPRO Orange. Subsequently, the midpoint of this unfolding transition is derived from the observed melt curve to yield the melting temperature ($T_m$) of the protein.

The identification of an optimum-solubility buffer and the subsequent transfer of the protein into this buffer can lead to the production of crystals from crystallisation screens which had previously only led to extensive protein precipitation, as was demonstrated for the novel folded AF2059 protein from *Archaeoglobus fulgidus* which was known to precipitate in most buffers during protein concentration experiments [324]. Therefore, the identification of optimum-solubility buffers for XplA and XplA-HD may allow the proteins to be concentrated to the high concentrations (5-40 mg/mL) often used in crystallisation experiments while minimising precipitation.

4.4.1.1: *XplA-HD Buffer Screen*

XplA-HD was subjected to the JBS Solubility Kit buffer screen (Table 2.2, Materials and Methods), which consist of a set of 24 different buffers, each at a concentration of 50 mM and with a pH range from 4.5 to 9.0. Buffers from this screen include
commonly used KPi-, Tris- and HEPES-based systems and also some less common buffer systems using PIPPS and ammonium acetate. Relatively large variations in stability could be observed when varying the buffers (Figure 4.5), and a few buffers appear to be generally more favourable for protein stabilisation, in particular sodium/potassium phosphate (pH 7.0) (buffer 15), Bis-Tris (pH 6.0) (buffer 9) and imidazole (pH 8.0) (buffer 19) were the most stabilising buffers. The identification of an imidazole based buffer system being favourable for protein stabilisation is unsurprising, given that imidazole is an XplA ligand. Sodium/potassium phosphate (pH 7.0) gave the highest average $T_m$ value of 44.7 °C for XplA-HD, which was 3.4 °C higher than the next best buffer, 50 mM Bis-Tris (pH 6.0) with an average $T_m$ of 41.3 °C. A table of all the averaged $T_m$ values is shown in Table A2 in the addendum. The sodium/potassium phosphate (pH 7.0) buffer was subsequently used to further the trials to include the addition of additives to improve protein stability.

Figure 4.5: Thermofluor assay of XplA-HD buffer screen. Negative first derivatives of the melt curves produced by XplA-HD are shown in a range of different buffers from the JBS Solubility Kit. The minimum value represents the midpoint of the unfolding transition, providing a quick and robust method for determining the $T_m$ of samples in individual wells. Curves that shifted the minimum furthest to the right represent conditions in which the protein is more stable. A large variation in the $T_m$ values is observed, and only the buffers from which a $T_m$ could be derived are shown. dF/dT represents fluorescence change at different temperatures. The figure legend shows the buffer numbers from the solubility kit (Table 2.2 – section 2.2.24)
4.4.1.2: XplA-HD Additive Screen

The buffer identified in the previous experiment, 50 mM potassium phosphate (pH 7.0) was prepared, and protein solutions were prepared in this buffer and supplemented with additives from the JBS Solubility Kit, which consist of a set of 14 different additives including DTT, BME and NaCl. The whole complement is listed in Materials & Methods (Table 2.3). Figure 4.6 shows the results of the additive screen and showed that 80 mM NaCl (additive 3) (38.5 °C) improved the thermal stability of XplA-HD by 2.6 °C compared to XplA-HD in just 50 mM KPi (pH 7.0), which yields a $T_m$ of 35.9 °C. In addition, 40 mM NaCl (buffer 2) (37.4 °C) and 8% glycerol (buffer 5) (37.3 °C) also increased the $T_m$ of XplA-HD compared to additive-free buffer. Additives BME and DTT, which as shown in Chapter 3 (3.3.2.5), are capable of binding XplA also improved thermal stability compared to the control sample. As will be discussed in Chapter 5 (5.6.2), this is consistent with the observation that the full length enzyme is more stable in buffers containing moderate concentrations of NaCl. Furthermore, it is well known that glycerol, and to a lesser extent salt, have a stabilising effect on most proteins [327-329]. However, the use of glycerol as an additive in crystallisation buffers is undesirable, as at high concentrations its stabilising effect may be detrimental to crystal growth, a process of controlled protein aggregation. Salt in crystallisation buffers is also often avoided, as salt crystals can introduce a degree of crystal ambiguity. Although most experienced structural biologist are able to distinguish between salt and protein crystals. Moreover, the crystallogenesis of a cofactor-containing protein also negates this problem, as these proteins tend to have colour.
Figure 4.6: Thermofluor assay of XplA-HD additive screen. Negative first derivatives are shown of the melt curves produced by XplA-HD supplemented with additives from the JBS Solubility Kit. The derived $T_m$ values were compared to the $T_m$ of the protein in buffer alone (50 mM KPi (pH 7.0)) to determine the effect of the additives on protein stability. No melt curves were observed for some of the additives, likely due to unfavourable interactions which destabilised the protein, and only small variations of $T_m$ values were observed from those that did stabilise XplA-HD. The $T_m$ values are significantly lower than the $T_m$ of the optimal buffer observed in the buffer screen as it was not possible to reproduce the optimal buffer to the same quality of the commercially bought buffer. The figure legend shows the additive numbers from the additive kit (Table 2.3 – section 2.2.2.4).

4.4.1.3: Effects on XplA-HD Thermal Stability by Ligand Binding

The additive screen was repeated to determine the effect of ligand binding on the thermal stability of XplA-HD, and to test if a different additive may maximise protein stability when in a ligand-bound form compared to the ligand-free form. RDX and imidazole were chosen as the ligands to study given their relevance to crystallographic studies. The assays show that RDX binding improves the thermal stability of XplA-HD in all buffers which also yielded $T_m$ values from the ligand-free protein. The addition of 80 mM NaCl (additive 3) (41.0 °C) again maximised the $T_m$ of RDX-bound XplA-HD, and was 2.4 °C higher compared to the enzyme substrate complex in 50 mM KPi (pH 7.0) buffer alone (38.6 °C). This is similar to the stabilising effect of this additive on the ligand-free protein, which saw an increase of 2.6 °C when supplemented with 80 mM NaCl. Imidazole binding improved the
thermal stability of XplA-HD further, and consistent with imidazole being a stronger ligand to XplA than RDX. The addition of 80 mM NaCl increased the $T_m$ of the imidazole-bound protein to 42.4 °C, compared to 41.0 and 38.5 °C for RDX-bound and ligand-free proteins, respectively. This indicates a near 4 °C increase in protein stability when in the imidazole-bound form, compared to ligand-free. A similar stabilising effect of additive 3 was also observed in the imidazole-bound complex compared to that previously observed for the RDX-bound and ligand-free forms, with a 3.4 °C increase in stability compared to the control sample (39 °C).

The same samples and buffers were used in the ligand-free and ligand-bound experiments to reduce the degree of variability which may otherwise be introduced if different samples and/or buffers were used. The protein behaved in the same way in each respective additive condition, whether in the ligand-free or -bound form, e.g. additives which stabilised the ligand-free protein also stabilised the ligand-bound forms.
Figure 4.7: Thermofluor assay of ligand-bound XplA-HD additive screen. Negative first derivatives are shown of the melt curves produced by XplA-HD supplemented with additives from the JBS Solubility Kit and when bound to RDX (A) and imidazole (B) at concentrations of 1 and 0.5 mM, respectively. The derived $T_m$ values were compared to the $T_m$ values of ligand-free protein in the same additives. Ligand binding did not affect the stabilising or de-stabilising effect of individual additives, and the additives which were more favourable for stabilising the ligand-free protein were also more favourable for stabilising both ligand-bound forms. The figure legends show the additive numbers from the additive kit (Table 2.3 – section 2.2.24).

The analysis of XplA-HD using the Thermofluor assay identified 50 mM KPi buffer (pH 7.0) with 80 mM NaCl as the most stabilising buffer/additive composition.
Subsequently, all XplA-HD crystallisation samples were made up in this buffer, with the NaCl concentration increased to 100 mM. A table of all the averaged $T_m$ values of the additive screen with and without ligands is shown in Table A3 in the addendum.

### 4.4.1.4: XplA Buffer Screen

In addition to using the Thermofluor assay to determine the buffer conditions which maximise the stability of XplA, the assay was also used to identify conditions that encourage FMN retention. This was achieved by performing the experiment in the absence of SYPRO Orange, and by using the fluorescence of the FMN cofactor. In this setup the thermal unfolding process of the protein can be analysed by following the FMN fluorescence, which would increase as it is released from the unfolding protein. Subsequently, the conditions which were most conducive for FMN binding would be reflected in the buffers yielding the highest $T_m$ values. However, unlike the XplA-HD, the full length enzyme was not amenable to analysis using the standard Thermofluor assay as it had high initial fluorescence. This may possibly be due to its multi-domain nature and the possibility of the FMN cofactor dissociating prior to the start of the experiment. Therefore, the buffer/additive combination identified from the analysis of XplA-HD was also used for the intact protein. However, in light of the issues of FMN loss (Chapter 5; 5.6.1), XplA samples were also made up in 50 mM HEPES, 100 mM NaCl buffer (pH 7.0) for crystallisation experiments, as phosphate ions have been demonstrated to compete with the FMN phosphate group in binding the phosphate pocket [214].

### 4.4.2: Multiangle Laser Light Scattering

Multiangle laser light scattering (MALLS) is a technique that can be used to assess the potential crystallisability of a newly purified protein by providing preliminary analysis of the homogeneity of a protein. This is achieved by determining its polydispersity index, which can give an indication of the presence of multiple aggregation or oligomeric states. Proteins that exist as a single homogenous population containing particles of uniform size are more promising candidates for crystallisation than those in which the protein is polydisperse [330]. MALLS data can also be used to derive the $R_h$ and $M_r$ of the protein. MALLS was carried out on the full
Chapter 4: Results II

length XplA and XplA-HD to determine if the purified samples were monodisperse, and samples were prepared as described in Materials & Methods (2.2.23).

4.4.2.1: \textit{MALLS Analysis of XplA}

The MALLS experiment involved loading XplA onto a size-exclusion column and monitoring the flow-through for light scattering and absorbance at 280 nm. The trace from these detectors, shown in Figure 4.8A, shows a single peak eluting from the S200 column after \(~14\) mL, indicating a single population of protein. Zimm fitting of the data from this peak provided values of the average \(R_h\) at 14.3 nm and the \(M_r\) at 60.6 kDa demonstrating that XplA is a monomer under the conditions tested. This is consistent with the observation that most P450s exist as monomers in solution, e.g. all the \textit{M. tuberculosis} P450s so far characterised using MALLS have been shown to be monomeric in solution [303, 307, 331]. In contrast, the well characterised BM3 can exist in monomeric and dimeric states depending on the concentration, and experimental evidence suggests the enzyme is functional in the dimeric state [332]. However, BM3 being an entirely self-sufficient enzyme, does not need to interact with other redox partners and requires only the pyridine nucleotide coenzyme, whereas P450 and redox partner interactions are paramount to the activity of most P450s, and dimerisation of the P450 could be a hindrance to these interactions.

A plot of the \(M_r\) values against elution volume (Figure 4.8B) shows a steady value of \(M_r\) over the duration of the peak. This indicates that the XplA sample is highly monodisperse, with a single globular protein species. The slight difference between the value for the \(M_r\) determined by MALLS and those observed for XplA predicted from the amino acid sequence (62 kDa) occurs mainly as the \(M_r\) reported by MALLS is inferred from the \(R_h\) of the sample. The MALLS data collected for XplA demonstrate that the protein has been isolated in a form highly suitable for crystallographic studies.
Figure 4.8: **MALLS analysis of XplA.** (A) Trace from a MALLS run of a sample of XplA (5 μM). Light scattering is indicated as a voltage and refractive index is a relative scale. These values are used to calculate the \( M_r \). A single peak elutes between 13-15 mL indicating only one species of XplA. The peak between 7-9 mL is known as the void and is expected in this experiment. (B) A plot of the molecular masses against the elution volume from the same run. The \( M_r \) of XplA derived from Zimm fitting was 60.6 kDa.

### 4.4.2.2: MALLS Analysis of XplA-HD

The XplA-HD was also analysed using MALLS, and the trace from a sample of XplA-HD eluting from the size-exclusion column shows a single peak eluting after ~16 mL (Figure 4.9A), indicating a single population of protein. Zimm fitting of the data from the single peak provided values of the average \( R_h \) at 8.0 nm and the \( M_r \) at 46.7 kDa, this indicates that the XplA-HD sample is monomeric with a single globular protein species. The \( M_r \) derived from MALLS is close to the predicted molecular weight of XplA-HD (45.7 kDa) determined from its amino acid sequence. The larger volume required to elute XplA-HD, and smaller \( R_h \) compared to XplA is consistent with XplA-HD being a smaller protein. A plot of the \( M_r \) values against elution volume (Figure 4.9B) shows a slight variation of \( M_r \) over the duration of the peak. However, the variation in the \( M_r \) values did not increase the polydispersity index, and the sample was shown to be monodisperse.
Figure 4.9: MALLS analysis of XplA-HD. (A) Trace from a MALLS run of a sample of XplA-HD (5 μM). Light scattering is indicated as a voltage and refractive index is a relative scale. A single peak elutes between 15-17 mL indicating only one species of XplA-HD. (B) A plot of the molecular masses against the elution volume from the same run. The average $M_r$ of XplA derived from Zimm fitting was 46.7 kDa.

4.4.3: Circular Dichroism Spectroscopy

CD spectroscopy investigates the differences in the absorption of left handed polarised light versus right-handed polarised light, which arise due to asymmetry in the structure of a molecule, i.e. amino acids in the case of proteins. CD spectroscopy is frequently used as a tool to investigate the structure of a protein in the absence of a solution or crystal structure provided by NMR or X-ray crystallography, respectively. In particular, CD spectroscopy is frequently used to examine the secondary structural integrity of a protein, whose spectral features arise in the FUV region of the spectrum (190-250 nm), due to the peptide bonds. The peptide bonds produce CD signals when they are located in a regular folded environment, and the α helices, β sheets and other random coil structures each contribute to the CD spectrum with their own characteristic shape and magnitude [333, 334]. The approximate fractions of each secondary structure can be estimated by analysing the FUV CD spectrum of the molecule and determining the fractional amounts of each by comparison to reference spectra. An absence of regular structure would result in zero CD intensity, while an ordered structure results in a spectrum which can contain both positive and negative signals. Normally, an α-helix gives a strong CD signal with minima at approximately 208 and 222 nm, whereas a β-sheet gives a rather weaker intensity spectrum with a single minimum at ~217 nm [335-337].
The tertiary structure can be analysed in the NUV spectral region (260-320 nm), since at these wavelengths the aromatic amino acids and disulphide bonds give rise to broad signals throughout the NUV spectrum. In this region, the CD signals are predominantly attributable to phenylalanine, tyrosine and tryptophan residues, and disulphide bonds only contribute weak signals [338]. For the P450s, the CD spectra in the visible region (320-600 nm) normally report on heme environment, and can be used to identify the redox state of the heme iron and the presence of ligands.

4.4.3.1: CD Analysis of XplA

The FUV CD spectrum of XplA is similar to the CD spectra reported for well characterised P450s, and indicative of a protein containing considerable α-helix secondary structural content as expected of a P450 [290, 339]. The FUV CD spectrum of XplA displays a weak positive maximum at 199 nm, and minima at 210.3 and 221.3 nm (Figure 4.10A). The negative CD signal at 221.3 nm is more intense compared to the minimum at 210.3, which is reversed compared to CYP121 and may indicate more contribution from β-sheets in XplA [290]. The FUV CD spectrum was analysed using the K2D3 program, which estimated the secondary structural content of XplA to be composed of 64.69% α-helix and 9.44% β-sheet. Based on studies of other P450s, the amount of α-helix is typically >50% [340], indicating that the secondary structural of XplA is quite similar to other P450s. The presence of the small flavodoxin domain, which has similar amounts of α-helix and β-sheet secondary structure content, does not significantly perturb the major P450 FUV CD spectrum. The secondary structural content of XplA, as predicted by K2D3, is significantly different to the secondary structural content of BM3. Munro et al. reported that the contents of α-helix and β-sheet in each case were 34% and 30%, respectively [341]. However, the secondary structure content was predicted using the CONTIN procedure [342]. Furthermore, in addition to the heme- and FMN-binding domains, BM3 also contains a relatively large FAD-binding domain.

CD spectra in the NUV region are characteristic of individual P450 isoforms. The NUV (Figure 4.10B) spectrum of XplA displays a broad positive CD signal between 275-308 nm, with peaks at 285.4 and 291.2 nm, respectively. In the visible CD spectrum (320–600 nm; Figure 4.10C), there are spectral features relating mainly to the chiral signal from the heme cofactor. Negative CD signals arise in the region between 309-419 nm, with minima located at 353.6 nm, close to the electronic absorption maximum at 353
nm corresponding to the heme δ peak, and at 406.4 nm near the position of the Soret maximum in the electronic absorption spectrum. A weak positive CD signal is also observed at 429.6 nm. The visible CD spectrum is similar in overall shape to those reported for other P450s, e.g. CYP121 and CYP126 [290, 303]. However, it is significantly different to the NUV CD spectra of the heme domain of BM3, which does not display a positive CD signal at ~425 nm, and instead has a strong negative CD signal near 410 nm [290]. This is reversed compared to XplA and CYP121 where the feature at 350 nm has a more intense negative CD signature.

Figure 4.10: CD spectra of XplA. (A) The FUV CD spectrum of XplA is shown at a concentration of 2.0 μM. The protein was in 10 mM KPi (pH 7.5). The FUV CD spectrum of XplA exhibited minima at 221.3 and 210.3 nm, and a weak maximum at ~199 nm. The secondary structural composition was calculated as described in the Materials & Methods (2.2.21). (B) NUV CD spectrum recorded using 20 μM XplA. (C) UV/Visible CD spectrum recorded using 20 μM XplA.
4.5: X-ray Crystallographic Studies of XplA

4.5.1: Crystallisation Trials of XplA

XplA was purified to homogeneity using Ni-NTA affinity and size-exclusion chromatography as described in Chapter 3 (3.2). MALLS also demonstrated that the protein was monodisperse and exists as a single species in solution, properties which are conducive for crystal growth. Furthermore, buffer conditions which improved the retention of the FMN cofactor were identified, and will be discussed in Chapter 5 (5.6.2).

Initial crystallisation trials of the full length protein were performed across a range of conditions in 96-well plates using the sitting drop method. A range of protein concentrations from 5-25 mg/mL were screened with a number of commercially available crystallisation screens from Molecular Dimensions (Newmarket, UK). Co-crystallisation experiments were also performed in order to improve the chances of obtaining crystals, with RDX, given its relevance as the only reported substrate, and also with imidazole as it significantly improves the thermal stability of the enzyme. The structures of proteins are typically solved in a ligand-bound form, because ligand binding normally increases the stability of the protein by locking the protein in a more rigid conformation [343-346]. To ensure saturation of enzyme with ligand, a concentration of ligand equal to the protein concentration plus 10 times the $K_d$ was used in co-crystallisation experiments. Substrate binding was confirmed prior to setting up crystal trays by using thin walled cuvettes (0.1-0.5 mm) to confirm the spin state of XplA. Solid substrate was also added to preparations of XplA to avoid the use of DMSO, and samples were prepared by addition of solid RDX to the protein and left stirring O/N at 4 °C. Undissolved substrate was removed by centrifugation and the supernatant was checked for ligation state as above, and showed that XplA was in the fully HS state, confirming binding of solubilised substrate.

Crystals were only obtained from the Pact Premier HT-96 (PP) crystallisation screen, in trays containing 25 mg/mL of protein with 1 mM RDX and incubation at R/T. Specifically, only three conditions were suitable for crystal growth, they were B7, B8 and B9. All three conditions share a common buffer system consisting of 100 mM MES, (pH 6.0) 20% PEG 6K, and 200 mM NaCl, 200 mM NH$_4$Cl and 200 mM LiCl for B7, B8 and B9, respectively. Crystals obtained from the trials by the sitting drop
method were too small for X-ray diffraction experiments but were judged to be of good quality, and were unique and of regular ordering (Figure 4.11). Although the colour of the crystals indicate the substrate is unlikely to be bound, as crystals of substrate-bound P450s in the HS state normally display brown-like colour, whereas crystals of XplA were red in colour.

![Figure 4.11: Crystallogenesis of XplA.](image)

**Figure 4.11: Crystallogenesis of XplA.** Crystals obtained from co-crystallisation trials of XplA with RDX. Only three conditions from the PP crystallisation screen yielded crystals. Specifically these were B7-B9, which shared a 100 mM MES, (pH 6.0) 20% PEG 6K buffer system, with 200 mM NaCl, 200 mM NH₄Cl, and 200 mM LiCl for B7, B8 and B9, respectively. Shown are the crystals obtained from B7 (A) and B9 (B). Crystallisation conditions were 100 nL of 25 mg/mL protein solution with 1 mM RDX, 100 nL of mother liquor and incubation at R/T. Crystals appeared O/N, at which stage they had reached their maximum size.

The three buffers from which crystals were obtained were reproduced and the crystallisation experiments scaled up using 400 nL of the protein solution plus 400 nL of the mother liquor using the conditions described above. Numerous crystals of sufficient size and quality were obtained (Figure 4.12A). However these crystals deteriorated before they could be frozen (Figure 4.12B), and hitherto the conditions that yielded these crystals have not yet been reproduced. In previous studies, the XplA-HD structure reported by Sabbadin et al. was solved using crystals grown using the full length protein, demonstrating the tendency of XplA to proteolyse, and crystallise as a cleaved protein containing only the P450 domain [246].
Chapter 4: Results II

Figure 4.12: Formation of diffraction quality XplA crystals. Crystals were obtained by co-crystallisation of XplA with RDX. The crystals shown were grown in 100 mM MES, (pH 6.0) 20% PEG 6K, and 200 mM NaCl (A). The crystals deteriorated ~24 h after first being observed (B). Crystallisation conditions were 400 nL of 25 mg/mL protein solution with 1 mM RDX and 400 nL of mother liquor. The tray was left at R/T, and crystals appeared O/N, at which stage they had reached their maximum size.

4.6: X-ray Crystallographic Studies of XplA-HD

4.6.1: Crystallisation Trials of XplA-HD

Initial crystallographic trials of XplA-HD were done as with the full length protein in 96-well plates and using the same commercially available screens. The trials focussed on the ligand-free and RDX-bound forms of the protein, as an imidazole-bound structure of the XplA-HD has already been reported [246]. Crystals were obtained from both forms and from multiple screens and conditions, with the RDX-bound form crystallising more readily than the ligand-free form. In particular, conditions from the Morpheus screen were very conducive to crystal growth with at least half the conditions producing crystals for the ligand-free form, and even more for the RDX-bound form. This is likely due to the fact that a third of the conditions in the Morpheus screen use a buffer system containing imidazole, and therefore likely crystallised in the imidazole-bound state. However, many of the wells which contained crystals did not use an imidazole buffer. The majority of the conditions produced a large number of small crystals, but some produced large isolated crystals, which are typically more promising for X-ray diffraction experiments (Figure 4.13). As with the crystals of the full length enzyme, the red colour of the crystals indicate RDX is unlikely to be bound. The colour of the crystals were also pale compared to
the deep red crystals normally observed for P450s, and may be due to a high solvent content in the crystals. All crystals obtained were of the same hexagonal shape and colour for both forms in all but one of the screens. Crystals obtained from the PP screen at 4 °C, using 25 mg/mL protein with 1 mM RDX were thin layered crystals of a deeper red colour which stratified. These crystals and other small crystals were used in micro-seeding experiments to encourage nucleation from these crystal seeds and in attempts to deter the further growth of small crystals.

Crystals were obtained from a range of protein concentrations from 12.5 mg/mL to 40 mg/mL, both at R/T and at 4 °C. However, crystal growth was slower at 4 °C, with crystals typically appearing after one week. Some had already grown to their final size at this stage, whilst others reached their final size typically one week after appearing. Most of the crystals grown at R/T appeared as quickly as O/N, at which point they had reached their maximum sizes. Furthermore, crystals grew more readily at R/T, although there were a lot more conditions that yielded large numbers of smaller unusable crystals. Regardless of incubation period or temperature, the same crystal forms were obtained.

Figure 4.13: Crystallogenesis of ligand-bound XplA-HD. Crystals of RDX-bound XplA-HD obtained from the Morpheus crystallisation screen are shown. Two types of crystals were obtained. Most of the wells contained numerous small crystals which had a very pale red colour (A), while other wells produced a few larger crystals (B). Crystallisation conditions were 100 nL of 25 mg/mL protein with 1 mM RDX, 100 nL of mother liquor and incubation at R/T. Crystals appeared O/N, at which stage they had reached their maximum size.

To obtain sufficiently large crystals for X-ray diffraction experiments, the crystallographic trials were scaled up to 400 nL drops of both the crystal solution and
mother liquor. The most promising screens from the initial trials were selected for further optimisation. These included the PP, Clear Strategy Screen (CSS) 2 and Morpheus screens. Crystals were again obtained, and were of sufficient size and quality in terms of regularity for X-ray diffraction experiments (Figure 4.14). Crystals obtained from the PP screen formed layers, as with the initial trial, and could not be separated, and those that grew in isolation were too thin to be used. Several crystals were taken from the Morpheus screen, including crystals from A2 (100 mM mix of imidazole; sodium cacodylate; MES; Bis-Tris (pH 6.5), and 60 mM mix of MgCl2; CaCl2, and 30% ethylene glycol-PEG 8K), B6 (100 mM sodium HEPES; MOPS (pH 7.5), 90 mM NaF; NaBr; NaI, and 30% glycerol-PEG 4K) and G5 (100 mM sodium HEPES; MOPS (pH 7.5), 100 mM Na-formate; NH4-acetate; Na3-citrate; NaK-tartrate (racemic); Na-oxamate, and 30% PEGMME 550-PEG 20K) for X-ray diffraction experiments.

**Figure 4.14:** Formation of diffraction quality XplA-HD crystals. Crystals of RDX-bound XplA-HD obtained from the Morpheus crystallisation screen. Crystallisation conditions were 400 nL of 25 mg/mL protein solution, 400 nL of mother liquor. The tray was left at R/T, and crystals appeared O/N, at which stage they had reached their maximum size. Crystals shown are those from wells A2 (A) containing; 100 mM mix of imidazole; sodium cacodylate; MES; and Bis-Tris (pH 6.5), 60 mM mix of MgCl2; CaCl2, and 30% ethylene glycol-PEG 8K, and G5 (B) containing; 100 mM sodium HEPES; MOPS (pH 7.5), 100 mM Na-formate; NH4-acetate; Na3-citrate; NaK-tartrate (racemic); Na-oxamate, and 30% PEGMME 550-PEG 20K. Relative composition of the buffer mixes unknown.

### 4.6.2: Structure of Imidazole-bound XplA-HD

Crystals selected for X-ray diffraction analysis were frozen by Professor David Leys, and data collection was done by Dr Colin Levy (both University of Manchester), and at the Diamond synchrotron (Didcot, Oxfordshire). Crystallisation conditions from the
Chapter 4: Results II

Morpheus screen were cryo-ready, thus addition of cryoprotectant was unnecessary. The structure was determined using molecular replacement using the published structure of imidazole-bound XplA-HD (2WIY) [246]. Molecular replacement was done using Phaser from the CCP4 program suite (Version 6.2.0) [280, 347, 348]. Structure refinement was done using Refmac5, also from CCP4 [281, 282]. The structure was analysed using Coot version, and was refined following 14 reiterations [349]. The data collection and refinement statistics are shown in Table 4.1.

Table 4.1: X-ray data collection and refinement statistics for R. rhodochrous strain 11Y XplA-HD.

<table>
<thead>
<tr>
<th>#Space group</th>
<th>I4₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell (Å)</td>
<td>a=b=136.1 c=75.1</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>43-2.30 (2.36-2.30)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>6.9 (38.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Reflections</td>
<td>28983 (2142)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>6.3 (3.2)</td>
</tr>
<tr>
<td>Rwork</td>
<td>20.1 (28.0)</td>
</tr>
<tr>
<td>Rfree</td>
<td>24.2 (30.0)</td>
</tr>
<tr>
<td>Average B (Å²)</td>
<td>28.0</td>
</tr>
<tr>
<td>R.m.s.d.</td>
<td></td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>0.023</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>1.885</td>
</tr>
<tr>
<td>Ramachandran</td>
<td></td>
</tr>
<tr>
<td>preferred (%)</td>
<td>94.86</td>
</tr>
<tr>
<td>allowed (%)</td>
<td>4.63</td>
</tr>
</tbody>
</table>

The XplA-HD was crystallised in a different space group from previously published structures, I4₁ compared to P2₁ and P2₁2₁2₂ for 2WIV and 2WIY, respectively [246]. A comparison with the previously solved 2WIY structure reveals little overall change in the structure (r.m.s.d. of 0.518 Å over 392 Ca atoms), with the exception of the residues comprising the BC loop region (220-241) (Figure 4.15). This loop has undergone a minor rearrangement as a consequence of the binding of a PEG molecule from the mother liquor at the active site region. Although imidazole was completely removed from XplA-HD prior to crystallisation, and despite the fact that 1 mM RDX substrate was used during crystallisation, the only conditions in which XplA-HD crystallised were those containing imidazole as a component of the crystallisation buffer, such as buffers from the Morpheus crystallisation screen, or from samples still
contaminated with imidazole. Electron density from the crystals clearly indicates that an imidazole is bound to the XplA heme iron (Figure 4.16C).

![Figure 4.15: Structure of the imidazole-coordinated XplA heme domain. Overlay of the XplA heme domain-PEG-imidazole ternary complex (green) with the previously determined XplA-imidazole complex structure (grey) (PDB code: 2WIY). The region most affected by PEG binding, the BC loop region, is coloured in blue for the ternary complex. PEG and imidazole ligands are shown in atom coloured sticks.]

The PEG extends into the heme binding pocket and is in van der Waals contact with a range of hydrophobic residues from the BC loop (W224, W230 and L238). In addition, hydrogen bonding contacts are observed with Q325 and T236, the latter mediated via water molecule W4. The bound PEG is also in van der Waals contact with 6th ligand to the heme iron, imidazole (Figure 4.16A). In addition to van der Waals contacts with several I helix residues (V391, M394 and A395), a water mediated hydrogen bonding network is made between imidazole, W1, W2 and P437-Q438 (Figure 4.16B).
Figure 4.16: Stereo view of the ligand-protein interactions observed in the XplA-HD ternary complex. (A) Selected residues contacting PEG are shown as sticks. The residues which are displaced due to the presence of the PEG are shown in their original positions in the previously solved structure in grey lines (2WIY). (B) Selected residues contacting imidazole are shown as sticks with a ribbon representation for the I helix region (residues 388-398). (C) Electron density map showing the ternary complex, the PEG is shown in ball and stick model.

The unusually high affinity of XplA for imidazole can be understood from the strong complementarity between the XplA distal heme binding pocket and the ligand (Figure 4.17). It was suggested previously that the I helix conformation relative to the heme iron is key to determining the affinity of azole-type ligands for the heme iron [302]. In the case of XplA, the I helix is oriented such that imidazole coordinates the heme iron with ideal geometry. This is due not only to the severe kink in the I helix as it passes over the heme plane, but also to the fact that residue 395 is an Ala as opposed to the usual Thr/Ser residues observed in the majority of P450s at this position. This avoids significant clash between the imidazole and the hydroxyl side chain that is usually present at this position. The imidazole ligand is furthermore bound within a hydrophobic pocket with high shape complementarity made by residues V391, M394...
and A395. In addition to the Fe-imidazole ligation, an extended water-mediated hydrogen bonding network is established with the other imidazole nitrogen atom. The unusually tight binding of imidazole to XplA can thus be explained from the geometry and chemical complementarity of the XplA active site.

In contrast, other bulkier azole ligands bind much less tightly or not at all to XplA (e.g. histidine, 1-phenylimidazole and fluconazole). This can similarly be understood from the strong complementarity between ligand and active site observed in the imidazole-XplA complex. Any substituents on the imidazole ring would cause considerable van der Waals clashes with V391, M394 and A395.

![Figure 4.17: Comparison of inhibitor binding in CYP121 and XplA-HD.](image)

(A) The kink in the XplA I helix introduced due to the absence of the highly conserved acid/alcohol pair, allows the imidazole to bind directly above the heme iron (green). The I helix of CYP121 approaches closer to the heme, which reduces the space for ligand binding and forcing ligands to adopt a bent conformation as observed with the fluconazole-bound complex (grey). (B) Steric hindrance due to S237 in CYP121 prevents the imidazole group of fluconazole from sitting directly above the heme iron (grey). The respective position in XplA, is occupied by A395 which does not intrude into the ligand binding pocket (green).

### 4.6.3: Refinement of Crystallisation Conditions

In light of the issues with imidazole and PEG ligation found in all crystal structures determined, binding experiments were performed to establish if XplA-HD binds either fatty acids, which are known substrates of physiological relevance to P450s, or the PEG molecules that were used as precipitation agents. Optical titrations were done with a range of saturated fatty acids including hexanoic (C6) and lauric (C12) acids. However, minimal spectral changes were observed for all fatty acids tested.
Chapter 4: Results II

(Figure 4.18). To exclude the possibility of the fatty acids even entering the heme active site and occupying a position which does not directly perturb the heme iron, as demonstrated by the reverse type I substrate TNB, the titration was repeated in the presence of sub-saturating concentrations of imidazole (10 μM). The titrations with hexanoic and lauric acids were repeated and spectral changes at the imidazole-ligated Soret peak (424.5 nm) monitored. If a fatty acid molecule approached the near to the imidazole, it might be expected that the XplA spectrum would be perturbed, e.g. the fatty acids might shift the equilibrium more in favour of the imidazole-bound complex by sitting on top of the imidazole and forcing it remain in the bound to the heme. This would be reflected as an increase of the 424.5 nm peak corresponding to the LS imidazole-bound complex. Alternatively, it could compete with the imidazole and prevent its binding, pushing the spectrum back towards that typical for the LS ligand-free form. However, as before, minimal spectral changes were observed during the fatty acids titrations (Figures not shown).

**Figure 4.18: Optical titrations of XplA-HD with fatty acids.** UV/Visible absorption spectra of XplA-HD (~3 μM) titrated with hexanoic acid (A) and lauric acid (B). Only the initial (black) and final (red) spectra from the titrations are shown, corresponding to the enzyme in the ferric ligand-free state, and following addition of 0.5 mM hexanoic or lauric acid, respectively. Spectral shifts suggestive of fatty acid binding to the XplA-HD active site (in either LS or HS directions) are not observed.

The titrations with short chain fatty acids demonstrated that XplA does not interact strongly with fatty acid substrates. This supports the observation that the ‘sausage’-like molecule protruding from the heme active site in the XplA-HD structure is a PEG molecule. Furthermore, for the ‘sausage’-like molecule to be a fatty acid, the bound fatty acid must have been retained from the *E. coli* cell lysate throughout the entire
purification procedure or acquired through contamination of buffers or columns. P450s are known to occasionally co-purify with endogenous *E. coli* compounds including fatty acids and phospholipids [350]. However, given that XplA shows no real affinity for fatty acid substrates, this is unlikely to be the case with XplA.

Optical titrations with commonly used PEG molecules in crystallisation experiments were done as with the fatty acids. Should any of the PEGs tested bind to XplA, these could be excluded from the crystallisation experiments. However, as with titrations with fatty acids, optical titrations with a range of PEGs including PEG 1K and PEG 2K MME induced no spectral changes, also indicating that PEGs do not bind strongly to XplA-HD, or at least do not interact measurably with the heme itself (Figure 4.19). To investigate if these PEGs could at least enter the heme active site, imidazole was also titrated with XplA supplemented with PEG 1K or PEG 2K MME. The opposite experiment was done with the fatty acids where imidazole was pre-bound instead to the heme, and titrated with fatty acids. The saturation of the heme active site with PEG would likely perturb the imidazole binding affinity, due to competitive binding. However, the $K_d$ of imidazole binding was not significantly altered, binding with a $K_d$ of 2.40 ± 0.06 μM in the presence of PEG 1K (500 μM) and a $K_d$ of 2.05 ± 0.06 μM in the presence of PEG 2K MME (500 μM). The $K_d$ of imidazole binding to the native enzyme had been determined to be 1.57 ± 0.04 μM in Chapter 3 (3.3.2.3). This suggests these PEGs do not enter the active site of XplA-HD. Owing to the constrained nature of the XplA active site, it is unsurprising that neither fatty acids nor PEG bound to XplA under the conditions used. However, it is clear from the crystal structure determined that, under certain crystallisation conditions, a PEG molecule enters the XplA active site substrate access channel.
Figure 4.19: XplA-HD interactions with polyethylene glycols. UV/Visible absorption spectra of XplA (~3 µM) bound to PEG 1K (A) and PEG 2K MME (B) at concentrations of 500 µM, and titrated with imidazole to a final concentration of 12 µM. Spectral changes were identical to previous titrations of XplA with imidazole in the absence of PEGs. Only the ligand-free enzyme (black), PEG-bound (red) and final spectra from the titrations with imidazole are shown. The insets show binding curves derived and fitted using Equation 2.1, as before for the imidazole titration to the native enzyme, generating $K_d$'s of $2.40 \pm 0.06$ µM with PEG 1K and $2.05 \pm 0.06$ µM with PEG 2K MME.

Given the propensity of XplA-HD to crystallise preferentially with imidazole, it was decided to try and soak crystals in RDX in the hope of replacing imidazole with RDX. This could easily be monitored by a change in colour of the crystals if RDX binding does occur. To accomplish this, individual crystals were soaked with mother liquor supplemented with 5 mM RDX (although it was not fully dissolved). However, no colour change was observed within the time scale expected (within seconds of soaking). These crystals were left soaking for a week at 4 °C with RDX-containing buffer, but no optical changes were observed and it is thus likely that the crystals remained in the imidazole-bound configuration.

4.6.4: Alternative Purification Protocol for XplA-HD

To resolve the persistent issue of imidazole co-crystallisation, alternative methods were sought to purify XplA-HD in order to avoid imidazole contact altogether. With the knowledge that bulky azole compounds did not ligate to the heme, the use of histidine was explored as an alternative to imidazole for elution from a Ni-NTA affinity chromatography column. Histidine, being a much larger compound than
imidazole, is unlikely to ligate to the heme of XplA-HD, despite containing an imidazole ring side chain.

Optical titrations of XplA-HD with histidine showed minimal spectral changes, demonstrating that histidine does not ligate to XplA-HD (Figure 4.20A). In addition, EPR spectra of ligand-free and histidine-bound XplA-HD were also recorded (Figure 4.20B), showing there is no difference between the two samples, and confirming that histidine does not ligate to XplA-HD. Interestingly, the EPR spectrum of XplA-HD purified in the absence of imidazole also detected the presence of the same two LS heme species described earlier, suggesting the LS species with $g_z = 2.56$ is not due an imidazole-ligated species, and supports the explanation that both LS species are likely due to the 6th water ligand binding in altered conformations. Subsequently, XplA-HD was eluted from Ni-NTA columns using 100 mM histidine as described in Materials & Methods (2.2.7). The purity of the eluted protein was comparable to protein eluted using imidazole (gels not shown). XplA-HD was passed down a gel filtration column as a final polishing step and crystallisation trials were performed as before. However, no crystals were obtained from ligand-free or RDX-bound samples, possibly due to insufficient purity of the protein, or since the enzyme can only crystallise in the imidazole-bound form. However, this is unlikely as many conditions in the Morpheus crystallisation screen use an imidazole based buffer system.

**Figure 4.20: XplA-HD and histidine interactions.** (A) UV/Visible absorption spectra of XplA-HD (3 μM) titrated with histidine (1 M) to a final concentration of 10 mM. Minimal spectral changes are induced indicating that histidine likely does not interact with the heme iron. The inset shows the initial (black) and final (red) spectra from the titration. (B) EPR spectra of ligand-free XplA-HD (200 μM) (black) and XplA-HD (200 μM) with 10 mM histidine (red). As with the EPR signature of the intact XplA, two LS species are detected with $g_z$ values 2.56 and 2.49 in ligand-free XplA-HD. The introduction of histidine does not appear to affect either of the two species present in the ligand-free sample.
Following the inability to obtain crystals from protein purified using histidine, an alternative purification protocol was developed in an attempt to improve protein purity. XplA-HD was ultimately purified using DEAE ion exchange and hydroxyapatite column chromatography, and the protein polished using a gel filtration column (Figure 4.21). At this stage XplA-HD was the major band and a sample was also taken at this point for EPR analysis to determine the ligation state of the heme iron. The EPR spectrum mirrors those attained from ligand-free samples purified using Ni-NTA affinity chromatography purification (Figure 4.22), confirming that there are two LS species in the EPR spectrum of ligand-free XplA-HD. Therefore, it is unlikely that the duality of signals is caused by imidazole as was first thought, but likely an internal ligand perturbing the water-ligated heme species which is conformationally trapped in the cryogenically frozen state.

**Figure 4.21: Hydroxyapatite and size-exclusion chromatography purification of XplA-HD.** (A) 10% (w/v) SDS-PAGE gel showing XplA-HD fractions eluted from a hydroxyapatite column. The gel shows from left to right; molecular weight marker (1) and fractions containing XplA-HD (2-7). (B) 10% (w/v) SDS-PAGE gel showing from left to right; XplA-HD following gel filtration (8) and molecular weight marker (9).
Figure 4.22: EPR spectrum of hydroxyapatite and size-exclusion chromatography-purified XplA-HD. The EPR spectrum of XplA-HD purified in the absence of imidazole again shows the presence of the same two LS species present in all ligand-free XplA samples analysed so far. This indicates the species with \( g_z = 2.56 \) (initially assigned to the imidazole-ligated species) may not be due to imidazole binding. However, there is an increase in relative intensity of the \( g_z = 2.56 \) species and a corresponding decrease of the \( g_z = 2.49 \) species, compared to previously analysed samples.

4.6.5: Crystallisation Trials of XplA-HD Purified Using the Alternative Protocol

Protein purified using the method described in 4.8.5 was used to set up crystallisation trials using Molecular Dimension crystallisation kits JCSG+, Morpheus, PP and CSS I and II at various concentrations of protein ranging from 10-25 mg/mL, at 4 °C, using 100 nL drops. Crystals were obtained from various conditions in different screens. Importantly, these crystals were of a different morphology to those previously obtained, which were rhombic bipyramidal (Figure 4.14 – section 4.6.1), and were cubic with the centres “pinched in” mildly (Figure 4.23). Furthermore, these crystals were of a colour consistent with a P450 in the substrate-bound HS state, and appeared more brown than red in colour when visualised through the microscope. Subsequently, the crystallisation experiments were scaled up using 200 nL drops of both the protein solution and of the mother liquor, and using Morpheus and PP screens to obtain larger crystals. The most promising crystals of diffraction quality
were obtained from the Morpheus screen, and were of the same shape and colour to the crystals obtained in the initial trials using 100 nL drops. Several of these crystals were selected for x-ray diffraction experiments.

**Figure 4.23: Co-crystallisation of RDX-bound XplA-HD.** Crystals shown were obtained from crystallisation trials of DEAE ion exchange, hydroxyapatite and size-exclusion chromatography-purified XplA-HD in the RDX-bound form. Crystals shown are those obtained from the PP screen, specifically B8 (A) containing; 100 mM MES (pH 6.0), 200 mM NH₄Cl and 20% (w/v) PEG 6K, and D7 (B) containing; 100 mM Tris (pH 8.0), 200 mM NaCl and 20% (w/v) PEG 6K. Crystallisation conditions were 200 nL of 25 mg/mL protein solution with 1 mM RDX, 200 nL of mother liquor and incubation at 4 °C. Crystals appeared after a few days, and after one week reached their maximum size.

Unfortunately, these crystals did not diffract to the resolution achieved using previously obtained crystals, and a useful structure could not be obtained. However, crude analysis of the data sets did identify a ligand bound to the heme iron. The electron density of this molecule was too large for a water molecule, but too small for RDX. Therefore, given the previous observation from UV/Visible spectroscopic data that XplA binds a RDX breakdown product, this molecule could potentially be a breakdown product of RDX. However, it is unclear at this point how the RDX molecule is degraded. RDX, being a recalcitrant environmental pollutant, should be a relatively robust compound, and there are no endogenous sources of electrons provided which could be used by XplA to reduce RDX. A possible explanation is that the RDX is reduced by the X-ray beam when the crystals are mounted for X-ray diffraction experiments, or that RDX is reduced by heme iron that is itself reduced in the beam. However, this is a contentious issue and it is presently unclear how this occurs, or if indeed the X-rays are responsible for RDX reduction.
4.7: Summary

Biophysical characterisation of XplA using MALLS showed that the purified protein was a homogenous monomeric protein. CD spectroscopic analysis of XplA in the FUV region showed the protein had considerable α-helix and minor β-sheet secondary structure content, which is consistent with P450s that typically contain >50% α-helices. Features in the UV/Visible region of the CD spectra were similar to those previously reported for CYP121 and CYP51B1, with negative CD signals close to the wavelengths of the heme Soret and δ-peaks in the absorption spectra, respectively. MALLS analysis of XplA-HD also showed that the separated P450 exist as a homogenous protein consisting of a single species in solution. Complementary analysis of the heme domain using the Thermofluor assays demonstrated that the protein was stabilised in 50 mM sodium/potassium phosphate, with 80 mM NaCl (pH 7.0), and crystallisation experiments were done in a similar buffer containing 100 mM NaCl buffer. Crystallisation experiments for the full length protein were performed in a HEPES based buffer, as phosphate ions are known to interfere with FMN binding by competing with the FMN phosphate group for the phosphate binding pocket. The biophysical analysis of purified XplA and XplA-HD showed that both proteins were suitable for crystallisation screening in their isolated states.

Crystallisation screening of the intact XplA identified only three conditions that were conducive for crystal growth. The three conditions all shared a MES based buffer system supplemented with NaCl, LiCl, and NH₄Cl, respectively. Scaling up of the crystallisation experiment in these buffers resulted in larger crystals more suitable for X-ray diffraction experiment. However, these crystals deteriorated prior to freezing, making them unsuitable for X-ray diffraction experiments. Up to this point, the conditions that led to the production of crystals from the intact protein have yet to be reproduced, and it remains unclear why crystals readily formed from that particular sample. Furthermore, the colour of the crystals, which were rather red for a substrate-bound P450, suggests they were of the native enzyme and RDX was not bound. The inability to crystallise the intact protein is likely due to a combination of XplA being a multi-domain protein and the tendency of the protein to lose its FMN cofactor, which subsequently introduces another level of conformational flexibility.

In contrast to the intact protein, the XplA-HD crystallised very readily, including in co-crystallisation experiments involving a variety of ligands. However, it later became
Chapter 4: Results II

apparent that the protein crystallised only when bound to imidazole despite the presence of saturating concentrations of other ligands. As described in Chapter 3 (3.3.2), imidazole binds very tightly to XplA with an apparent $K_d$ of $1.57 \pm 0.04 \mu M$. Imidazole was thought to be fully removed by extensive dialysis, and this was confirmed by spectral analysis through demonstrating the appearance of the Soret peak at $\sim 417$ nm. However, this could not be confirmed with EPR due to the heterogeneity of heme ligation state species always associated with XplA when analysed by EPR. Furthermore, the binding of other ligands was confirmed prior to crystallogenesis by analysis of the position of the XplA-HD Soret peak (these crude spectra were not shown). It is likely that, although UV/Visible spectroscopy showed that the heme was predominantly in the substrate-bound form when performing co-crystallisation experiments with RDX, or in the ligand-free form when crystallising the ligand-free protein, a small population of imidazole-bound species remains despite the exhaustive attempts to remove this species. The UV/Visible absorption spectrum of this minor imidazole-bound species is likely masked by the major ligand-free or substrate-bound species, and subsequently it is the former species which crystallises during crystallogenesis experiments despite its relatively low concentration as compared to the other major heme population.

The structure determined using crystals obtained from co-crystallisation experiments with RDX showed that, rather than the RDX molecule, an imidazole molecule was bound to the heme in its place. This finding was consistent to the previously published structure of XplA-HD [246]. The structure showed that the imidazole ligand is able to bind the heme iron in an ideal geometry, due to the severe kink in the I helix as it passes over the heme, and the absence of the conserved Ser/Thr I helix residue which would otherwise have introduced a clash between the side chain hydroxyl group and the imidazole. Furthermore, imidazole is bound within a hydrophobic pocket with high shape complementarity consisting of residues from the I helix. In addition to the Fe-imidazole ligation, an extended water-mediated hydrogen bonding network is established with the other imidazole nitrogen atom. In addition to the bound imidazole, a PEG molecule was also shown to extend from the heme binding pocket and into the substrate access channel. The PEG makes van der Waals contact with the imidazole and with residues lining the hydrophobic access channel. The introduction of the PEG into the structure introduces conformational changes of residues near the BC loop to accommodate the PEG. Despite the
appearance of PEG in the structure, it was later shown that the PEG molecule only interacts with XplA-HD under crystallisation conditions, as optical titrations with a range of PEGs and fatty acids did not provide evidence of interactions in the heme environment.

Through further crystallisation screening of XplA-HD using protein purified without Ni-NTA affinity chromatography columns, i.e. avoiding imidazole altogether, crystals were also obtained from co-crystallisation experiments with RDX. Promisingly, the crystals did not share the same shape and morphology to those previously obtained, and the colour of these crystals was consistent with the crystals of a P450 in the substrate-bound HS form. However, X-ray diffraction analysis of these crystals showed they did not diffract well when compared to those previously obtained, although from the data collected it was determined that a ligand was bound to the heme, and that this ligand was not imidazole. The electron density of this ligand was not consistent with RDX or water, suggesting that it is a likely RDX degradation product. However, it is unclear how RDX was degraded as no obvious sources of electrons are present in the crystallisation buffers.

Although no structure has yet been determined using selected crystals obtained in this manner, this crystal form offers promise as the crystals are clearly of the protein in the substrate-bound form prior to X-ray diffraction. Furthermore, these crystals demonstrate that XplA-HD can be crystallised in the absence of imidazole, paving the way for structural characterisation of XplA-HD in the ligand-free or other ligand-bound forms. The identification of at least one of the RDX degradation products ligating the heme facilitates co-crystallisation experiments with these compounds. Both the aerobic- and anaerobic-dependent breakdown products, NDAB and MEDINA are possible ligand candidates as they are highly similar compounds containing nitrogen atoms that can ligate to the heme iron [106]. The determination of substrate- and product-bound structures of XplA-HD would be of great assistance in determining the catalytic mechanism of RDX breakdown. With further optimisation of the purification and crystallisation procedures, more promising crystals may yet be obtained.
Chapter 5: Expression, Purification & Characterisation of the Flavoproteins of the XplA/XplB Redox System
Chapter 5: Results III

5.1: Introduction

The vast majority of cytochromes P450 enzymes catalyse their reactions with electrons derived ultimately from a pyridine nucleotide coenzyme, and utilising redox partners, which typically consists of a flavodoxin/ferredoxin reductase and a small flavodoxin or ferredoxin protein. The best characterised of these redox partners, as discussed in Chapter 1, are the FAD- and FMN-dependent CPR from mammalian microsomes and the putidaredoxin reductase/putidaredoxin system from *P. putida*. CPR is unique amongst P450 redox partners as it is a diflavin reductase incorporating both redox partners as a single enzyme and containing both FAD and FMN cofactors. Like most other P450 redox partners, CPR can reduce several different P450s, and is responsible for supporting the activity of all the microsomal P450s [183]. CPR is also unique because, together with the cindoxin reductase/cindoxin P450 redox system from *C. braakii* and BM3, represent the only such systems which utilise both FAD and FMN cofactors naturally [9, 183, 194, 206, 220]. Most bacterial P450s utilise redox partners consisting of a FAD-binding reductase and a Fe-S cluster protein.

As more P450s are identified from newly deciphered genomes of novel organisms, so too will new redox partners be identified. These cognate redox partners are typically, but not exclusively, found adjacent to the P450 in the genome. The implication of a P450 enzyme being involved in RDX degradation, and later identified as XplA, also led to the discovery of its reductase partner, encoded by *xplB*. XplB is a NADPH-dependent flavodoxin reductase containing a molecule of FAD as its cofactor and shares sequence homology with AdR [247]. The two genes are located adjacent to each other in the *R. rhodochrous* genome, and together form a novel P450 redox system which also utilises both FAD and FMN cofactors, with the FMN-binding flavodoxin protein already attached to the P450. The XplB enzyme has only been partially characterised and shown to be able to support XplA activity [106]. However, its biochemical and structural properties have yet to be defined.

In order to study the interactions between a P450-flavodoxin fusion enzyme and its redox partner, *xplB* from *R. rhodochrous* was expressed, and the protein purified and characterised. Furthermore, in light of the inability to crystallise the full length XplA, which would reveal information on the heme and FMN environments, the FMN-binding domain of XplA was expressed and purified independently of the P450
domain. This chapter focuses on the studies of the flavoproteins, XplB and the XplA flavodoxin, of this novel P450 redox system.

5.2: Molecular Biology of XplB

The xplB gene was synthesised by GenScript with E. coli codon optimisation and cloned into pET15b to produce the pET15b(xplB) construct. The entire sequence is displayed in Figure A2 of the addendum. This construct was transformed into tetracycline-resistant NovaBlue competent cells, from which fresh plasmid DNA was prepared and the cells used to maintain a glycerol stock to be stored at -80 °C.

Jackson et al. reported XplB to be a highly insoluble protein [106]. Therefore, to improve the chances of producing soluble XplB, expression trials were extended by cloning the xplB gene into other expression vectors. A non His6-tagged construct (pET11a(xplB)) was prepared by excising the xplB gene from pET15b(xplB) and cloning into pET11a using NdeI and BamHI restriction sites. Jackson et al. were eventually able to produce soluble XplB with ~50% FAD incorporation using a pEGX2T vector, with a GST-tag fused to the N-terminus. Cleavage of the GST-tag from the GST-XplB fusion protein led to the aggregation of the isolated XplB. However, close analysis of the UV/Visible absorption spectrum of the purified protein suggests the FAD may not be bound [106]. Furthermore, no other biochemical data were presented for the purified XplB.

The xplB gene was also cloned into the pCold TF vector (Takara Bio Inc). The pCold TF vector is specifically designed to improve the solubility of overexpressed recombinant proteins that have been difficult to isolate in a soluble state. The vector encodes cold shock proteins that up-regulate target protein production at lowered incubation temperatures. The temperature drop also suppresses production of other cellular proteins and temporarily halts overall cell growth. Recombinant proteins are synthesised bound to the C-terminus of a 48 kDa trigger factor (TF) chaperone protein with multiple cleavable sites in the linker region connecting the recombinant and TF protein to allow easy separation. To generate the pCold TF(xplB) construct, the xplB gene was excised from pET15b(xplB) using Ndel and BamHI restriction enzymes, and pCold TF was cut using the same enzymes. The resulting digests were run on a 0.8% agarose gel to separate inserts from vectors (Figure 5.1A). The xplB
gene and pCold TF vector were excised, gel purified and ligated as described in Materials & Methods (2.2.1/2.2.3). Successful ligation was confirmed by sequencing plasmids from successful transformants, followed by restriction digestion using NdeI and BamHI, and visualisation of a DNA fragment of the expected size for xplB (1,278 bp) in a DNA gel (Figure 5.1B).

![Figure 5.1: Cloning of pCold TF(xplB).](image)

(A) 0.8% (w/v) DNA gel showing the NdeI and BamHI double restriction digestion of pET15b(xplB) and pCold TF, and the separation of the excised xplB gene from the pET15b vector. The gel shows; DNA ladder (1), NdeI and BamHI digested pET15b(xplB) (2) and NdeI and BamHI digested pCold TF (3). (B) 0.8% (w/v) DNA gel showing NdeI and BamHI digested pCold TF(xplB), confirming the successful ligation of the excised vector and gene from the gel shown in (A). The lanes show; DNA ladder (4) and NdeI and BamHI digested plasmid (5).

5.3: XplB Expression Trials

The pET15b(xplB) construct was transformed into E. coli strains BL21 (DE3), HMS174 (DE3), Origami B (DE3) and Rosetta 2 (DE3); and pET11a(xplB) was transformed into BL21 (DE3) and HMS174 (DE3). Expression trials were conducted in those vector/strain combinations as described in Materials & Methods (2.2.5). The trials show that xplB was expressed in all strains tested. SDS-PAGE gels show a band with a size approximating to that of XplB (48 kDa) appearing in samples from cells where xplB expression was induced using IPTG (Figure 5.2A). Furthermore, xplB expression was notably improved in cells carrying the pET11a(xplB) construct (Figure 5.2B). Basal expression of xplB from the pET15b vector was very low. However, xplB was expressed to higher levels in absence of IPTG induction when cloned in the pET11a vector.
Figure 5.2: Time-lapse production of *R. rhodochrous* XplB. (A) 10% (w/v) SDS-PAGE gel showing the expression profile of XplB in HMS174 (DE3) cells carrying the pET15b(xplB) construct. A band corresponding to the expected size of the His<sub>6</sub>-tagged XplB (~48 kDa) indicated by the marked arrow appears in lanes where xplB expression was induced using IPTG. The gel shows samples (1 mL) taken at several time points from cells induced with 100 µM IPTG, and from un-induced cells as the control. The gel shows from left to right; molecular weight marker (1), time = 0 (2), 1 h control (3), 1 h induced (4), 2 h control (5), 2 h induced (6), 4 h control (7), 4 h induced (8), 24 h control (9) and 24 h induced (10). (B) 10% (w/v) SDS-PAGE gel showing the expression profile of XplB in HMS174 (DE3) cells carrying the pET11a(xplB) construct. Expression of xplB was induced as with the pET15b(xplB) construct, and the gel layout is as with gel (A). An overexpressed band corresponding to the expected size of untagged XplB (~45.8 kDa) is seen in lanes containing samples where xplB expression was induced using IPTG and indicated by the marked arrow.

Although overexpression of xplB was achieved, very limited amounts of the protein was purified from the soluble fractions recovered from cells carrying the pET15b(xplB) or pET11a(xplB) constructs following Ni-NTA affinity (Figure 5.3A) and DEAE ion exchange (Figure 5.3B) chromatography purification, respectively. Moreover, the protein was almost entirely insoluble.
As with XplA, HMS174 (DE3) cells produced the largest quantity of recombinant protein. Subsequently, this strain was selected for further optimisation to improve the production of soluble XplB. However, in light of the insolubility of XplB, other strains in which overexpression of XplB was observed were also used in the solubility trials, e.g. BL21 (DE3) cells. This involved varying growth conditions such as the
temperature, IPTG concentration, induction period, growth medium and addition of riboflavin. However, all conditions tested to improve soluble protein yield were unsuccessful.

The pCold TF(xplB) construct was transformed into BL21 (DE3) cells. The use of the TF protein to solubilise recombinant proteins proved successful with XplB and an overexpressed band corresponding to the size of the TF-XplB fusion protein (~100 kDa) was observed on a SDS-PAGE gel following Ni-NTA affinity chromatography purification of the soluble cell lysate (Figure 5.4). However, the protein was colourless, and not yellow as expected of a flavoprotein, indicating the absence of the FAD cofactor. This could be due to incorrect folding resulting in low FAD incorporation during protein synthesis or to FAD lost following cell breakage. Due to the lack of a yellow-like colour observed during purification, the former explanation is more likely. The co-expression of the N-terminal 48 kDa TF protein could be a hindrance to FAD incorporation during XplB synthesis, if the FAD cofactor is incorporated after the proteins have fully folded.

![Figure 5.4: Ni-NTA affinity chromatography purification of TF-XplB.](image)

10% (w/v) SDS-PAGE gel showing TF-XplB purification using Ni-NTA affinity chromatography. The gel shows from left to right; molecular weight marker (1), soluble fraction of cell lysate (2), eluate from a 10 mM imidazole wash (3), eluate from a 30 mM imidazole wash (4), 1 mL fractions from a 300 mM imidazole wash (5-9) and insoluble fraction (10). Expression conditions were; 100 µM IPTG, 20 °C, 24 h using pCold TF(xplB) in HMS174 (DE3) cells.

Attempts were made to separate XplB from the TF protein by cleavage using thrombin. Following small scale digests to optimise the digestion efficiency according
Chapter 5: Results III
to the manufacturer’s specification, the TF-XplB fusion protein was digested using a ratio of 0.02 U of enzyme per 10 μg of protein in 5 mL reaction volumes. However, cleavage was incomplete in the scaled up digests and the cleaved XplB could not be separated from TF using Ni-NTA affinity chromatography. Following thrombin cleavage, the TF protein which carries the N-terminal His6-tag is immobilised on the Ni-NTA column, whilst the separated XplB does not bind and is collected in the flow-through. However, SDS-PAGE shows that although >50% of the TF-XplB fusion protein is cleaved, the majority of the cleaved XplB co-elutes with the TF protein (Figure 5.5). Due to the inability to efficiently separate XplB from TF and the fact that FAD incorporation was not observed, expression of xplB using the pCold TF vector was not pursued further.

![Ni-NTA affinity chromatography purification of thrombin treated TF-XplB](image)

**Figure 5.5: Ni-NTA affinity chromatography purification of thrombin treated TF-XplB.** 10% (w/v) SDS-PAGE gel showing the purification of His6-tagged TF using Ni-NTA affinity chromatography following cleavage from the TF-XplB fusion protein using thrombin. The TF-XplB fusion protein was digested using a ratio of 0.02 U of enzyme per 10 μg of protein in 5 mL reaction volumes. The gel shows from left to right; molecular weight marker (1), flow-through (2), eluate from a 10 mM imidazole wash (3) and eluate from a 300 mM imidazole wash (4), with the appearance of several major bands corresponding to the expected sizes of His6-tagged TF-XplB (~100 kDa), His6-tagged TF (~52 kDa) and XplB (~45.8 kDa), respectively.

Although soluble XplB was not recovered from expression trials using pET15b and pET11a vectors, expression trials were scaled up using 1 L cultures of HMS174 (DE3) cells carrying the pET15b(xplB) construct. The cells were grown in TB medium and xplB expression induced with 100 μM IPTG at 20 °C for 24 h to determine if small quantities of XplB could be recovered from the soluble fraction. Scaling up produced
soluble XplB at a level less than that of some constitutively produced proteins in the 
*E. coli* cells, and it was not possible to separate XplB from these proteins using Ni-NTA 
affinity chromatography. Subsequently, pET15b(xplB) was transformed into NiCo21 
(DE3) (NEB) cells in order to minimise *E. coli* protein contamination of Ni-NTA 
affinity chromatography fractions. In NiCo21 (DE3) cells, glucosamine synthetase 
and three other endogenous *E. coli* proteins that bind to Ni-NTA columns have either 
been mutated to eliminate binding to Ni-NTA resin or tagged to enable rapid removal 
by chitin affinity chromatography [351]. However, the production of soluble XplB 
was inferior to that achieved using HMS174 (DE3) cells, and expression of *xplB* using 
this strain was not pursued further.

5.4: Expression & Purification of XplB

Ultimately, with further tweaking of the expression conditions using HMS174 (DE3), 
overexpression of soluble XplB was achieved. It was determined that reducing the 
volume of growth medium from 1 L to 0.5 L in a 2 L Erlenmeyer flask greatly 
improved the solubility of XplB, and that this was likely dependent on aeration 
conditions. The overexpressed band observed on SDS-PAGE gel with an approximate 
size of ~48 kDa was confirmed to be XplB by mass spectrometry (Figure 5.6A). The 
purity of XplB was greatly improved following purification using hydroxyapatite 
chromatography, and two major bands were observed on a SDS-PAGE gel of XplB-
containing fractions (Figure 5.6B). The fractions were initially identified by analysing 
the fractions with UV/Visible spectroscopy to determine protein containing fractions, 
as determined by their 280 nm peak. Both bands were confirmed by mass 
spectrometry to be XplB, and the purest fractions were pooled and further purified 
using size-exclusion chromatography.

Unfortunately, despite soluble XplB being overexpressed, as was observed for the TF-
XplB fusion protein, the isolated XplB was colourless indicating the absence of the 
FAD cofactor. Furthermore, it was noticed that during the first purification step, and 
specifically during the first wash step after binding the protein to a Ni-NTA column, 
the flow-through was unusually yellow. This indicated that either a flavoprotein or 
free flavin was eluting. The flow-through was analysed by UV/Visible spectroscopy 
and was revealed to be free flavin, likely FAD. This was confirmed by a SDS-PAGE,
which showed the absence of an overexpressed band corresponding to the expected size of XplB. Furthermore, free and bound flavins typically display slightly altered spectra, characterised by a shoulder at ~470 nm when the FAD is bound to a protein [176, 352-354]. The purification of recombinant flavoproteins which display sub-stoichiometric incorporation of the flavin cofactor has previously been reported, e.g. FNR from *Streptomyces griseus* purifies with ~50% FAD incorporated [355]. Moreover, FMN loss is more commonly observed, e.g. FMN loss is also observed during purification of the *E. coli* flavodoxin protein WrbA [356]. Therefore, it is quite unusual for a FAD-binding flavoprotein to completely lose its FAD. This suggests that the purified XplB has very weak FAD affinity or that it is incorrectly folded, leading weak FAD affinity.

Figure 5.6: Ni-NTA affinity and hydroxyapatite chromatography purification of XplB. (A) 10% (w/v) SDS-PAGE gel showing XplB purification using Ni-NTA affinity chromatography. The gel shows from left to right; molecular weight marker (1) and eluate from a 100 mM imidazole wash (2) with an overexpressed band corresponding to the expected size of His-tagged XplB (~48 kDa) indicated by the marked arrow. Expression conditions were; 100 μM IPTG, 20 °C, 24 h using pET15b(xplB) in HMS174 (DE3) cells. (B) 10% (w/v) SDS-PAGE gel showing XplB fractions eluted from a hydroxyapatite column. The gel shows from left to right; molecular weight marker (3) and fractions containing XplB as the major band (4-12) indicated by the marked arrow.

In light of the problem of FAD loss observed during purification of XplB using Ni-NTA affinity chromatography, other chromatographic techniques including DEAE ion exchange chromatography were tested as the first step to purify XplB to determine if FAD would be better retained in other columns. Encouragingly, a large yellow band was observed bound on the DEAE column. However, following extensive washing the
yellow band could be seen slowly moving down the column, which was again confirmed to be free FAD as above. This shows that XplB is produced with FAD-bound, and that separation of XplB from the soluble cell lysate facilitates dissociation of FAD from XplB.

Despite the issue of inadequate FAD incorporation in the isolated protein, large amounts of soluble and sufficiently pure XplB was recovered following purification using Ni-NTA affinity and hydroxyapatite chromatography, and approximately 29 mg of protein was purified from 12 L of culture (~2.4 mg/L).

5.5: Characterisation of XplB

5.5.1: Sequence Alignment of XplB with Related Ferredoxin/Flavodoxin Reductases

The amino acid sequence of XplB was compared with some of the most related oxidoreductases by multiple sequence alignment in order to identify conserved residues important for FAD binding, and to determine if they are retained in XplB (Figure 5.7). The alignment shows that XplB is structurally related to the GR family, which includes FprA from *M. tuberculosis* and mitochondrial AdR. XplB shares high sequence homology with several uncharacterised putative oxidoreductases, including a putative FNR from *Arthrobacter globiformis* (sp. NBRC 12137), with which XplB shares 43% sequence identity. However, XplB contains only one of the four conserved sequence motifs belonging to the GR family identified by Dym & Eisenberg [169]. The xhxGxGxxGxxxhxhxhxhxhE(D) motif (where h is a hydrophobic residue and numbers indicate the gap between residues) located at the N-terminus is part of the Rossmann fold, is present in all members of the GR family, and is known as the dinucleotide-binding motif [357]. The conserved glycines in the motif allow for a tight turn of the main chain, which permits close contact of the main chain to the FAD pyrophosphate [358]. The conserved acidic residue, Glu/Asp, hydrogen bonds to the ribose 2'‐hydroxyl group of the adenosine moiety (D42 in the aligned XplB sequence) [174].

The other sequence motifs are only partially conserved in the GR family, including the T(S)xxxxxF(Y)hhGD(E) motif first identified from rubredoxin reductase [359]. The residues in this motif are found on the seventh β-strand of the FAD binding domain,
Chapter 5: Results III

and also interact with the FAD pyrophosphate. As this motif and the other two motifs are only partially conserved, their absence from XplB is not surprising, and suggests that XplB is structurally divergent from the well characterised members belonging to this family.

Figure 5.7: Multiple amino acid sequence alignment of the XplB with other flavodoxin/ferredoxin reductases. Alignment of XplB sequence with a selection of the most closely related reductase sequences was done using the ClustalW2 alignment tool via the European Bioinformatics Institute [360]. Proteins aligned are; putative FNR from Erythrobacter (sp. SD-21) (Ery), AdR from human mitochondria (AdR), FprA from M. tuberculosis (FprA), putative FNR from Streptomyces viridochromogenes (sp. DSM 40736) (Svir), cindoxin reductase from C. braakii (CinB), XplB from R. rhodochrous and a putative FNR from Arthrobacter globiformis (sp. NBRC 12137) (Aglo). Conserved residues (G17, G19, G22 and D42 for XplB) that form part of the sequence motif typical of proteins containing the Rossmann fold and flanking the FAD phosphate group are in black text and underlined with yellow highlighted background. These are absolutely conserved in all sequences aligned. Also conserved is the tryptophan, shown in bold red text with grey highlighted background, which is thought to be involved in binding FAD, modulating its thermodynamic properties and in coenzyme reactivity. The double arginine motif that is involved in NADPH binding is shown in bold red text with grey highlighted background, which is absolutely conserved in all sequences aligned. Also conserved is the double arginine motif that is involved in NADPH binding.

XplB does not share significant sequence homology with many of the well characterised bacterial FNRs, including E. coli FldR and related proteins. Therefore,
XplB does not contain any of the conserved motifs belonging to the FNR family [169]. Binding of the FAD cofactor in AdR-like enzymes is not well defined when compared to FAD binding in the FNRs, which typically consists of two aromatic residues that form aromatic stacking and hydrophobic interactions with isoalloxazine ring. Furthermore, a third aromatic residue is present in some FNRs that interact with the adenosine moiety [181]. The location of this residue is responsible for the conformation taken by the FAD. The structure of AdR shows that the FAD binds in an elongated conformation with the ring pointing towards the NADP⁺-domain and the adenine portion buried in the FAD domain. The si-face of the ring is covered by backbone residues, whilst the re-face is solvent exposed, which contrasts with the FNRs, where the si-face is solvent exposed [181, 183, 184].

Unlike the FNRs, aromatic stacking interactions do not appear to be important for FAD binding. Instead, hydrogen bonding interactions between backbone residues and FAD appears to be more important in the AdR-like enzymes, e.g. the conserved acidic residue of the first motif hydrogen bonds to the ribose 2'-hydroxyl group as described earlier. The corresponding residue in the solved structure of bovine AdR is E38 (E69 of the aligned human AdR sequence, and D42 in XplB) [174]. Furthermore, many of the residues identified from the bovine AdR structure which make hydrogen bonds to the FAD appear to be conserved in the aligned sequences, e.g. the O2 atom of the isoalloxazine ring is hydrogen bonded to the amide of I376 (I407 in the aligned human AdR sequence). This residue is replaced by a leucine (L341) in XplB, but conserved in all other sequences aligned. Hydrogen bonds are also formed between the isoalloxazine ring O4 and N5 atoms and a water molecule that is fixed by the side chains of H55 and D159 (H86/D190 of the aligned human AdR sequence, and H59/D160 in XplB). Both residues are absolutely conserved in the sequences aligned [174]. The structure of AdR also identified R124 (R155 in human AdR) in helix α5 being involved in stabilising the FAD pyrophosphate through a water molecule, however this residue is not retained in any of the sequences aligned. The discrepancies between the residue numbering from bovine AdR and aligned human AdR is because the sequence derived from the bovine AdR structure started from residue 33 (S33), which was denoted as residue 1 in that respective sequence [174].

Furthermore, the tryptophan residue identified from FprA (W359) from M. tuberculosis that is important for FAD binding, thermodynamics and coenzyme
reactivity is also absolutely conserved amongst the protein sequences aligned, demonstrating its importance in this class of flavoproteins (W333 in XplB) [361]. However, one of the arginines from the double arginine (RR) motif identified from the same protein (R199/R200 in FprA), whose side chains are involved in electrostatic binding with the NADPH 2'-phosphate group in AdR-like enzymes, is not conserved in XplB or cindoxin reductase [362, 363]. The second arginine is replaced by a serine in both XplB and cindoxin reductase (S199 and S198 for XplB and CinB, respectively). The absence of this arginine is likely to decrease the affinity of NADP(H) binding as the presence of the arginines negates the negative charge of the phosphate group. Interestingly cindoxin reductase, with which XplB shares 37% sequence identity, has also been reported to be an insoluble protein and was not purified in a functional form [194].

Due to the diversity of ferredoxin/flavodoxin reductases, identification of residues important for FAD binding outside of the conserved sequence motifs is difficult, and many of the hydrogen bonds are formed with backbone amides. The absence of one of these residues would provide a possible explanation for the weak binding of FAD observed in XplB. However, the sequence alignment shows the retention of important residues for FAD binding identified from the structure of bovine AdR in XplB, and given the tendency of XplB to express as an insoluble protein, it is more reasonable to assume that it is incorrectly folded in its E. coli-expressed state.

5.5.2: \textit{Reconstitution of XplB with FAD}

The expression and purification of soluble XplB was achieved as described above, with the FAD cofactor at least partially incorporated prior to separation from the cell lysate. However, it is unclear at this point whether the enzyme is functional prior to isolation from the soluble cell lysate. The presence of \textit{E. coli} flavoproteins that can mimic XplB activity, e.g. flavodoxin reductase, make it difficult to determine whether XplB is active at this stage. However, \textit{xplB} was expressed in a functional state in \textit{A. thaliana}, as demonstrated by the faster growth rate of transgenic plants expressing both \textit{xplA} and \textit{xplB} genes compared to those carrying only the \textit{xplA} gene when grown in soil supplemented with RDX [257].
In order to perform biochemical and structural characterisation of XplB, it was necessary to reconstitute XplB with FAD to regenerate the holoprotein, and to this end several methods were attempted. Initial reconstitution efforts were done by incubation of XplB with excess FAD in a variety of different buffers, including the 24 different buffers from the JBS Solubility Kit listed in Table 2.2 in the Materials & Methods (2.2.24), followed by separation of excess and unbound FAD using centrifugation or by gel filtration on a PD-10 column. Buffer conditions were changed to test reconstitution efficiency, such as varying pH, ionic strength, and buffer constitution. Furthermore, binding was also monitored using fluorescence and UV/Visible spectroscopy by titrating FAD with XplB apoprotein and detecting fluorescence or spectral changes of the FAD cofactor, respectively. However, in all experiments conducted, no FAD binding was observed either because XplB does not bind FAD in its purified state or since binding is exceptionally weak and not detectable by the methods utilised. Moreover, many flavoproteins are less stable when their prosthetic group is lost, and it is possible that the stability of XplB deteriorates following FAD loss during purification, making it less amenable to FAD reconstitution.

XplB was also subjected to analysis using the Thermofluor assay to determine optimal buffer conditions for thermal stability. Experiments were also done with FAD added. Binding of the FAD cofactor is likely to increase the stability of the holoprotein, and this would be reflected in a positive $T_m$ shift compared to the apoprotein. However, informative results were not obtained from the experiments, and XplB in its purified state was not amenable to analysis using the Thermofluor assay due to high initial fluorescence. This may be due to the fact that the conformation of XplB in the current isolated state has exposed hydrophobic areas, to which the SYPRO orange dye could bind, leading to increased initial fluorescence.

Due to the inability to establish buffer conditions that would facilitate FAD binding to XplB, reconstitution was attempted by immobilising XplB on a Ni-NTA column under denaturing conditions using the chaotropic agent guanidine hydrochloride. Low concentrations of chaotrope have been shown to partially unfold proteins by disrupting hydrogen bonds and other non-covalent forces [144, 364, 365], and this process was utilised to determine whether partial unfolding of XplB may expose the FAD binding site, which may be closed off to prevent FAD dissociation when the
protein is fully folded. FAD reconstitution was achieved by washing the immobilised and partially unfolded protein with a concentrated FAD solution. It has been shown in reconstitution experiments with other apoflavoproteins that FAD binding is a simple bimolecular process that may occur rapidly once the chaotrope is removed [366]. However, it was apparent that, upon removal of the chaotrope and subsequent elution of the immobilised protein, either reconstitution was unsuccessful, or that any bound FAD dissociated from the protein once the column was re-equilibrated into non-denaturing buffer conditions.

### 5.5.3: Determination of the FAD Dissociation Constant

#### 5.5.3.1: Isothermal Titration Calorimetry

Isolation of XplB holoprotein has thus far been unsuccessful, and it remains unclear whether this is because XplB binds FAD exceptionally weakly, or since binding is discouraged due to incorrect folding of the purified protein. However, it appears that FAD is incorporated during protein synthesis, as FAD is clearly lost following immobilisation of XplB on chromatography columns.

In order to determine the FAD dissociation constant, XplB and FAD interactions were studied using ITC. ITC is a biophysical technique used to study the thermodynamic properties of a biochemical reaction, and is commonly used to analyse the binding of small molecules to larger macromolecules, such as a ligand to an enzyme. ITC can directly measure the $K_d$ of a reaction and the stoichiometry, binding enthalpy and entropy of the reaction [367].

XplB was titrated with FAD and the resultant heat changes were monitored following each titration. A number of attempts were made to determine the $K_d$ of FAD binding using ITC, by varying the concentrations of both XplB and FAD. The results were consistent, and the heat changes induced by XplB and FAD interactions were weak and could not be distinguished from instrument noise (Figures not shown). These findings were consistent with previous results that showed that XplB does not bind its FAD cofactor with any measurable affinity.
5.5.4: *The State of Folding of Purified XplB*

5.5.4.1: *CD Analysis of XplB*

To confirm whether XplB was correctly folded, CD spectroscopy was used to confirm the presence of secondary structure elements. CD spectra were acquired in the presence and absence of FAD to detect any changes in structure, which would demonstrate interaction between the XplB apoprotein and FAD.

The FUV CD spectrum of XplB defines it as a protein rich in β-sheet secondary structural content and displays maximum negative ellipticity at ~227 nm (Figure 5.8A) [368, 369]. Moreover, the FUV CD spectrum of XplB does not share similarities to other characterised FNRs, e.g. the FUV CD spectrum of *M. tuberculosis* FprA displays an α-helix rich secondary structure with significant β-sheet contribution (42% α-helix and 26% β-sheet) [363]. Analysis of the FUV CD spectrum of XplB with the K2D3 program predicted the secondary structural content of XplB to be 3.9% α-helices and 36.5% β-sheets. This suggests that the purified XplB is incorrectly folded, although it retains considerable β-sheet secondary structure. Furthermore, addition of FAD did not alter the FUV CD spectrum of XplB, indicating negligible effect on the secondary structure, possibly due to the absence of XplB-FAD interactions.

The NUV (Figure 5.8B) and visible CD (Figure 5.8C) spectra provide a unique ‘fingerprint’ for the enzyme, with characteristics often more clearly distinguishable from other flavoproteins than those provided by UV/Visible absorption spectra, which are usually very similar in all oxidised flavoproteins [353, 354, 363]. However, no chiral signatures were detected in the visible CD spectrum of XplB, due to the absence of FAD, and the CD spectra in NUV and visible regions are unlikely to be representative of the fully functional enzyme, particularly in the visible region where a cofactor signal is absent.
Figure 5.8:  CD spectra of XplB.  (A) The FUV CD spectrum (190-260 nm) was recorded using 10.0 μM XplB.  The FUV CD spectrum exhibits negative ellipticity at ~227 nm and is indicative of a protein with considerable β-sheet secondary structure. The secondary structural content of XplB was predicted using the K2D3 prediction tool, which shows that XplB has a secondary structure content of 3.9% α-helices and 36.5% β-sheets.  (B) NUV CD spectrum recorded using 100 μM XplB.  (C) CD spectrum of XplB in the NUV and visible regions recorded using 100 μM XplB.  The CD spectra in these regions provide a unique fingerprint for the enzyme, and flavoprotein features in the visible region arise due to contributions from the flavin cofactor.

5.5.4.2: MALLS Analysis of XplB

MALLS was also used to determine the state of folding in XplB, and also to confirm its $M_r$ and oligomeric state.  The MALLS trace from the eluate from a sample of XplB shows a single peak eluting from a Sephadex S200 gel filtration column after ~15 mL, indicating a homogeneous population of XplB (Figure 5.9A).  Also shown for comparison is the trace from the eluate of a second experiment performed in buffer supplemented with FAD (500 μM) (Figure 5.9B).  The results show that FAD does not alter the elution profile of XplB, and thus does not appear to affect its oligomeric state significantly.
Figure 5.9: MALLS analysis of XplB.  (A) Trace from a MALLS run of a sample of XplB (5 μM). Light scattering (red) is indicated as a voltage and refractive index (blue) is a relative scale. These data are used to calculate the $M_r$. A single major peak elutes between 14-16 mL indicating only one predominant species of XplB. The peak between 7-9 mL is known as the void and is expected in this experiment. (B) Trace from a second MALLS experiment of XplB (5 μM) in buffer containing FAD (500 μM). A single peak also elutes at the same volume as in the FAD-free sample. (C) Plot of the molecular masses against the elution volume for XplB (blue) and XplB with FAD (magenta). The average $M_r$ of XplB derived from Zimm fitting was 41.2 kDa in absence of FAD, and 42.4 kDa with FAD.

Zimm fitting of the light scattering and refractive index data of the peak eluting between 14-16 mL provided values of the average $R_h$ at 2.6 nm and the $M_r$ at 41.2 kDa for XplB without FAD, and of 2.3 nm and 42.4 kDa for XplB with FAD. This is slightly smaller than the molecular weight of XplB predicted from its amino acid sequence (45.7 kDa), but expected due to possible differences which may arise due to the use of the $R_h$ to derive the molecular mass, and also because clipping of the protein was also observed. However, the data indicates that the XplB sample is monomeric and a plot of the $M_r$ values against elution volume (Figure 5.9C) shows a steady value of $M_r$ over the duration of the peak, indicating XplB is monodisperse. In addition, there was no
Chapter 5: Results III

A major difference between the samples with and without FAD, indicating that the cofactor does not alter oligomerisation state of XplB. Most likely because binding did not occur. The $R_h$ of XplB is smaller than the $R_h$’s reported for flavoproteins glutamate synthase from *Rhodobacter capsulatus* (E1F1), which contained two multimeric active species with radii of 6.18 and 7.32 nm, respectively [370], and a NADH oxidase from *Lactococcus lactis*, which has an estimated radius of 8.7 nm [371]. However, these flavoproteins exist in oligomeric states with molecular masses exceeding 100 kDa, whereas XplB is homomeric and half the size. In addition, glutamate synthase and NADH oxidase do not belong to the oxidoreductase family of flavoproteins, therefore the differences observed in $R_h$’s is unsurprising. Furthermore, the estimated $R_h$ of XplB is consistent with the minimum $R_h$ expected for a globular protein of ~50 kDa, which is 2.4 nm [372].

5.5.5: **Activity of Purified XplB**

Successful reconstitution of XplB with FAD to regenerate the intact holoprotein remains elusive. Although XplB-FAD interactions were not demonstrated, the activity of the purified enzyme was tested in the absence and presence of FAD. The activity of XplB was shown by reconstituting the XplA/XplB redox system and monitoring XplA reduction and RDX turnover. Productive XplB and pyridine nucleotide coenzyme interactions were shown by monitoring pyridine nucleotide coenzyme oxidation and FAD reduction.

5.5.5.1: **NADPH-dependent Reduction of XplB**

Previous studies of XplB showed that it could oxidise NADPH [106]. Furthermore, the authors also demonstrated that XplB shows higher specificity for NADPH than NADH. XplB and FAD interactions were demonstrated for the first time following purification of XplB, by reducing FAD in solution supplemented with XplB using NADPH (Figure 5.10). UV/Visible spectroscopy shows that the addition of XplB apoprotein to free FAD lowers the 450 nm peak and also increases the intensity of the peak at 375 nm, and on addition of NADPH the FAD is fully reduced to the HQ state after 1 min. This indicates that the FAD may only bind XplB transiently, but can be encouraged to remain bound if it is kept in the reduced state, which is highly unstable to oxidation when free in solution. The reduction of FAD was not observed in the absence of XplB,
indicating that the reductive reaction was not achieved by direct reaction of NADPH with the free FAD.

![Figure 5.10: NADPH-dependent XplB reduction.](image)

**Figure 5.10: NADPH-dependent XplB reduction.** UV/Visible absorption spectra of oxidised FAD (40 μM) (black), with characteristic absorption maxima at 375 and 450 nm, and on addition of XplB (80 μM) (red), the peak at 450 nm is slightly decreased and the peak at 375 nm increases. The absorption of the FAD is completely bleached on reduction with NADPH (200 μM) (blue) after 1 min, indicating the formation of the FAD HQ species. No such reduction of FAD occurs in the absence of the XplB protein. An extinction coefficient of $\varepsilon_{450} = 11.3$ mM$^{-1}$ cm$^{-1}$ was used to confirm the concentration of the free FAD, which was made from a 1 mM FAD stock solution [269].

### 5.5.5.2: Reconstitution of the XplA/XplB Redox System

XplB-XplA interactions were shown by using CO-trapping assays to demonstrate electron transfer from NADPH, through XplB and onto the XplA heme. The electron transfer was visualised by the appearance of the Soret peak at 446.5 nm corresponding to the absorption maximum of the Fe(II)CO complex. RDX turnover was also demonstrated, and this was monitored by observing NADPH oxidation at 340 nm, or collapse of the RDX-bound HS complex with absorption maximum at 396 nm. There is significant contribution to the Soret peak by the added FAD cofactor, with the underlying flavin contribution shifting the Soret peak from 417 nm to 419 nm, and with a large shoulder at ~460-470 nm developing due to the oxidised FAD (Figure 5.11; black). Addition of RDX induces a LS to HS transition of the heme with
the Soret peak shifting to 390 nm (Figure 5.11; blue). However, a large residual shoulder is visible at ~421 nm which was not present in the UV/Visible absorption spectra of XplA titrated with RDX in the absence of XplB and FAD, and is possibly due to contribution of the FAD cofactor or to insufficient RDX to induce a full spin state shift to HS. Addition of NADPH leads to the reduction of RDX and regeneration of a ferric LS heme species with absorption maximum at 424 nm (Figure 5.11; red). The unusually long wavelength of this LS species is likely due to a combination of the breakdown product-bound species described in Chapter 3, and the absorbance of the oxidised FAD. The RDX is fully reduced after 4 min, and after another 60 min both the FAD and heme appear fully reduced (Figure 5.11; magenta).

**Figure 5.11: FAD- and XplB-mediated XplA activity.** UV/Visible absorption spectra are shown for ferric LS XplA (5 μM), with XplB (25 μM) and FAD (25 μM) (black). The Soret peak is at 419 nm with a broad shoulder from 460-470 nm due to the FAD. Addition of RDX (100 μM) induces a LS to HS transition of the heme and the Soret peak shifts to 390 nm, with a residual peak at 419 nm (blue). Addition of NADPH (200 μM) fully reduces the RDX and regenerates a LS heme species after 4 min. The absorption maximum of this LS species is at 424 nm (red). After 60 min, the XplA heme and XplB FAD cofactors are almost fully reduced as demonstrated by bleaching of the heme and flavin absorption features (magenta).

The successful reconstitution of the XplA/XplB redox system demonstrates that XplB is able to bind FAD in a productive manner and drive XplA-dependent turnover of RDX. However, on re-oxidation of XplB by XplA it is likely that the FAD cofactor
dissociates from the protein. XplA activity is only supported in the presence of free FAD and XplB following addition of NADPH.

5.6: **Analysis of the FMN Moiety of XplA**

5.6.1: **Quantification of the FMN Content of XplA**

On inspection of the UV/Visible absorption spectrum of purified XplA, a distinct lack of a shoulder at ~450 nm was observed that is characteristic of P450-flavoprotein fusion enzymes and which corresponds to the absorption of the oxidised form of FMN [106]. To investigate further, the FMN was removed and quantified according to the methods of Aliverti [268]. The FMN was removed from a known quantity of XplA by heat treatment. The non-covalently bound FMN cofactor remained soluble and visible by spectral and fluorescence analysis following heat precipitation of protein and centrifugation to separate the insoluble protein (Figure 5.12). Comparison with FMN standards indicated that there was only 30% incorporation of FMN relative to the heme bound to XplA. This suggests that the FMN binding affinity is unexpectedly weak, at least by comparison with microbial flavodoxins which typically have $K_d$ values of $<1 \mu$M for their FMN cofactor [207, 354, 373]. More stringent monitoring of the protein during purification gave a clear indication that the FMN was being lost during purification, as a yellow band was eluted much later than the protein band. Efforts turned to reconstituting the protein with FMN.
5.6.2: Reconstitution of XplA with FMN

Successful reconstitution of XplA with FMN was achieved coincidentally during a protein concentration step using two Vivaspin concentrators in two different buffers. It was established that buffers containing salt encouraged FMN retention. The flow-through in the concentrator with salt buffer was clear, while the flow-through from the second concentrator containing salt-free buffer was yellow. Subsequently, FMN reconstitution was simply achieved by incubation of XplA with excess FMN in 50 mM HEPES buffer (pH 7.2) containing 150 mM NaCl, and leaving O/N at 4 °C. Unbound FMN was removed by passing the sample through a PD-10 column. The difference in FMN content following XplA reconstitution is shown in Figure 5.13, where the enzyme before FMN reconstitution shows low absorption in the region of 450-500 nm where oxidised flavin absorbs maximally. On extensive incubation with FMN and gel filtration as described above, a prominent shoulder appears between 450-500 nm due to FMN binding. In order to retain FMN in XplA samples, subsequent experiments were done in buffers containing 150 mM NaCl.
reconstituted protein was quantified as before, and a ratio of 0.9:1 (90%) was calculated, again relative to the heme content of XplA.

![Graph](image)

**Figure 5.13: Reconstitution of XplA with FMN.** UV/Visible absorption spectra for a FMN-free XplA (~5.0 μM) sample prior to FMN addition (black), and following full reconstitution with FMN (red). Flavin spectral contributions are clear in the latter spectrum across the range from ~350-500 nm, and particularly in regions associated with absorbance maxima of oxidised FMN at ~380 and ~470 nm.

5.6.3: **Determination of the FMN Dissociation Constant**

To establish whether XplA binding to FMN is not particularly strong when compared to other flavodoxins, a fluorescence titration experiment was conducted to determine the dissociation constant of FMN binding. The property of flavin fluorescence quenching on protein binding was used to this effect. Flavin fluorescence is quenched upon binding to a protein because some of the energy absorbed from light at 480 nm (the excitation wavelength used in this experiment) is lost to the surrounding protein environment by rotational energies [374]. Therefore, the energy of the re-emitted photon is lower compared to that of the photon absorbed. Titration of XplA apoprotein into a FMN solution quenched the flavin fluorescence (Figure 5.14A). The fluorescence intensity change at 524 nm was used to determine the binding affinity of FMN. By subtracting the 524 nm excitation reading made at each titration point from the original reading made in absence of added protein, and then plotting the fluorescence change against the concentration of protein added, a binding curve was
produced, from which a $K_d$ value of $\sim 1.09 \pm 0.14 \, \mu M$ was determined by data fitting using Equation 2.1 (Figure 5.14B). XplA binds FMN with an affinity significantly lower than those observed for other bacterial flavodoxins, which typically bind FMN in the nM range, e.g. the *B. subtilis* flavodoxins YukN and YukP bind FMN with $K_d$ values of $14.6 \pm 0.6 \, nM$ and $25.2 \pm 1.0 \, nM$, respectively [354]. The unusually high $K_d$ of FMN binding in XplA explains why the FMN is lost during purification.

![Figure 5.14: Analysis of FMN binding to XplA.](image)

(A) FMN fluorescence emission spectra from the titration of FMN (0.25 μM) with FMN-free XplA to a final protein concentration of $\sim 13 \, \mu M$. The fluorescence of FMN is quenched as the protein concentration is increased. (B) The change in fluorescence intensity at 524 nm is plotted against the concentration of apoprotein titrated and data fitted using Equation 2.1 to yield a $K_d$ value of 1.09 ± 0.14 μM.

The determination of the FMN dissociation constant was performed in 50 mM HEPES (pH 7.5). As shown earlier, XplA and FMN binding is encouraged in buffers containing a moderate concentration of NaCl ($\sim 150 \, mM$). This suggests that the ionic strength of the buffer influences the dissociation constant of FMN binding. To determine the effect of ionic strength on FMN $K_d$, the above titration experiment was repeated in 50 mM HEPES, 150 mM NaCl (pH 7.5). Quenching of the flavin fluorescence was observed as previously described for the titration in salt-free buffer (Figure 5.15A). The fluorescence intensity change at 524 nm was used to determine the binding affinity of FMN as above, and the $K_d$ was determined as before to yield a $K_d$ value of 1.33 ± 0.04 μM. This is comparable to the dissociation constant of FMN binding determined in salt-free buffer. Lostao *et al.* showed that ionic strength does not have a large effect on FMN complex stability in *Anabaena* (sp. PCC 7119) flavodoxin because although it weakens phosphate interactions, it strengthens isoolalloxazine
ring-protein hydrophobic interactions [375]. Therefore, the dynamics of XplA and FMN interactions remain unclear.

Figure 5.15: Analysis of FMN binding to XplA in NaCl-containing buffer. (A) FMN fluorescence emission spectra from the titration of FMN (0.25 μM) with FMN-free XplA to a final protein concentration of 4 μM. The induced fluorescence changes are consistent with those observed for the titration in salt-free buffer. (B) The change in fluorescence intensity at 524 nm is plotted against the concentration of apo-protein titrated and data fitted using Equation 2.1 to yield a $K_d$ value of 1.33 ± 0.04 μM.

5.6.4: **Sequence Alignment of the XplA Flavodoxin with Bacterial Flavodoxins**

Alignment of the amino acid sequence of the XplA flavodoxin domain (residues 1-148 of full length XplA) with some of the most related short chain microbial flavodoxins (Figure 5.17) shows the retention of important residues involved in interactions with the FMN isoalloxazine ring and other parts of the FMN structure. A pair of aromatic residues (Tyr59 and Tyr97 in XplA) is conserved in most flavodoxins, and these interact with the *re*-face (Y59) and *si*-face (Y97) of the isoalloxazine ring [246]. The first of these aromatic residues is substituted by a histidine in MioC type flavodoxins and tryptophan in *Mannheimia haemolytica* flavodoxin. Structural studies of the *D. vulgaris* and *D. vulgaris* flavodoxins indicates that a series of bonding interactions is made by residues between P loop residues S10-T15 and the negatively charged FMN phosphate group, with hydrogen bonds formed between phosphate oxygen and backbone amide or Thr/Ser side chain hydroxyl groups of S10, T11, T12, N14 and T15 [376, 377]. Only three of these potential hydrogen bonding residues are conserved in XplA (T10, T12 and N14), and XplA E11 replaces T11 in the *Desulfovibrio* flavodoxins. The introduction of a glutamate in XplA has obvious
potential for repulsive interactions with the phosphate, although XplA E11 is conserved in some other flavodoxins. Aside from the P loop region, other FMN-binding loop regions are the 50’s loop (S55-D64 in E. coli MioC, corresponding to S57-A66 in XplA) and the 90’s loop (S92-G100 in MioC, D94-N102 in XplA) [378]. XplA does not retain a PEDPAE motif in the 120’s loop region near the C-terminal of the flavodoxin and considered to be present only the MioC subfamily of flavodoxins. However, the two proline residues from this motif (P132 and P136) are conserved in XplA.

**Figure 5.16: Multiple amino acid sequence alignment of the XplA flavodoxin with other flavodoxins.** Alignment of the XplA flavodoxin sequence (XplA residues 1-148) with a selection of the most closely related flavodoxin sequences was done using the ClustalW2 alignment tool. Proteins aligned are Erwinia carotovora (sp. atroseptica) flavodoxin (Ecar), cindoxin from C. braakii (Cin), MioC from Sodalis glossinidius (strain morsitans), MioC from E. coli (Eco), flavodoxin from Mannheimia haemolytica serotype A2 strain (Mhae) and flavodoxin from D. vulgaris (Dvul).

Aromatic residues (Y59 and Y97 for XplA) flanking the FMN isoalloxazine ring are in black bold text with grey highlighted background for each flavodoxin. The P loop, 50’s and 90’s loop regions of XplA are underlined. Residues implicated in FMN phosphate hydrogen bonding interactions in the XplA P loop are in bold blue text with yellow highlighted background for each flavodoxin. The P loop, 50’s and 90’s loop regions are in bold red text. The other XplA hydrophobic residue in these regions (other than Y57 and Y97) is coloured light blue in bold (Y100). The two 120’s loop proline residues in XplA are in bold white text with blue highlighted background, as are the respective amino acids in the MioC-type flavodoxins from S. glossinidius and E. coli.
5.6.5: **Determination of the XplA FMN Reduction Potentials**

One of the key roles of flavodoxins is to substitute for ferredoxin proteins that act as electron shuttle proteins which carry electrons typically from a flavodoxin/ferredoxin reductase to a P450, for example putidaredoxin in the P450cam redox system [379]. In order for flavodoxins to perform this function, they must modulate the potentials of the bound FMN cofactor so that electron transfer is thermodynamically favourable. The midpoint reduction potentials of the OX/SQ and SQ/HQ couples of the FMN are influenced by their immediate environment, e.g. by forming favourable interactions with one redox state over another [155, 156, 274, 380, 381].

The redox titration of full length XplA fully incorporated with FMN highlighted spectral changes at wavelengths at 470, 550 and 630 nm due to reduction of the FMN moiety from the fully oxidised state to the fully reduced HQ state. At 550 and 630 nm, the absorbance increases on generation of the SQ state, and decreases upon the transformation of this species to the HQ state upon further reduction (Figure 5.17A). To determine the potentials of the FMN cofactor, the absorption change at 630 nm was plotted against the applied potential and the data fitted to a 2-electron Nernst equation to yield a midpoint reduction potential of -80 ± 5 mV for the OX/SQ and -172 ± 6 mV for SQ/HQ couples, respectively (Figure 5.17B). The potentials are very positive compared to other bacterial flavodoxins, e.g. the midpoint reduction potentials for the OX/SQ and SQ/HQ couples of *E. coli* flavodoxin are -254 and -433 mV, respectively, but comparable to the potentials of the redox couples of cindoxin from *C. braakii*, which are -137 and -218 mV for the OX/SQ and SQ/HQ couples at pH 7.5, respectively [194]. The absorption change at 630 nm was selected to determine the potentials of the FMN, because minimal spectral changes due to heme reduction are observed at this wavelength, as compared to ~456 nm, where heme spectral changes are induced upon reduction, and could influence the determination of the FMN reduction potentials.
Figure 5.17: Determination of reduction potentials for the FMN cofactor in XplA. (A) Selected spectral data from a spectroelectrochemical redox titration of FMN-bound XplA (~8 µM). Only spectra collected in the early phase of the reductive titration (from ~ +75 to ~ -200 mV) are shown and reflect absorbance changes predominantly due to reduction of the FMN. Arrows indicate directions of absorption change during the reductive phase of the titration. At ~470 nm (near the absorption maximum for oxidised FMN) there is a decrease in absorbance in the early phase, reflecting conversion of FMN from oxidised through SQ to HQ. In the region at ~550 nm there is an initial increase in absorption due to FMN SQ formation, followed by a decrease as the SQ is reduced to FMN HQ. (B) Fit of A_{630} data (following FMN SQ formation and its further reduction to HQ) versus potential fitted using a 2-electron Nernst equation [30], giving midpoint potentials for the FMN OX/SQ (E₁) and SQ/HQ (E₂) couples of -80 ± 5 mV and -172 ± 8 mV, respectively.
Previous studies of *E. coli* MioC also commented on its relatively small number of negatively charged residues in the 50's and 90's loop regions that are in proximity with the FMN, and how these acidic residues might modulate the FMN potentials. XplA has a similar number of acidic residues in these regions as MioC (in slightly altered positions), and such alterations in the electrostatic environment of the XplA flavin is a possible explanation for its non-standard FMN redox properties \[378\]. Similarly, XplA has fewer aromatic residues in these two loop regions compared to e.g. *E. coli* FldA (but a similar number to *E. coli* MioC), which may also influence flavin stability and redox properties \[202, 378\].

5.7: Molecular Biology of the XplA Flavodoxin

To generate a plasmid construct expressing the XplA flavodoxin domain (residues 1-141 of XplA), two stop codons (TAA and TAG) were engineered in the linker region separating the FMN and heme domains. To achieve this, oligonucleotide primers XplA-FMFN (5’- CG GTT GCT GAA TGG GCC **TAA** TTT GCC GAA GCT CTG -3’) and XplA-FMNR (5’- CA GAG CTT CGG CAA **ACT** ATT AGG CCC ATT CAG CAA CCG -3’) were used, where non-complementary nucleotides introduced are in bold and the new stop codons are underlined. The mutagenesis PCR experiment was done using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene), following the manufacturer’s protocol. The mutant pET15b(xplA-FMN) construct was cloned into NovaBlue cells to generate a glycerol stock and to prepare fresh plasmid. pET15b(xplA-FMN) was subsequently transformed into *E. coli* expression hosts for expression trials. The mutant xplA gene was also excised from the pET15b vector and cloned into pET11a using Ndel and BamHI restriction enzymes to produce an untagged construct.

5.8: XplA Flavodoxin Expression Trials

Both constructs were transformed into BL21 (DE3), HMS174 (DE3), Origami B (DE3) and Rosetta 2 (DE3) cells, and expression trials were done in those cell strains and conducted as with xplB. The trials revealed that XplA flavodoxin was not expressed in any of the strains tested, as no overexpressed band was observed on SDS-PAGE gels (Figure 5.18). Furthermore, all bands of a size close to the approximate size of XplA
flavodoxin were identified using mass spectrometry, and all identifications were negative. Although the SDS-PAGE gels did not provide evidence of XplA flavodoxin production, they did show that the expression profile of the *E. coli* cells was altered upon the addition of IPTG. The production of *E. coli* proteins is significantly reduced when IPTG is used to up-regulate the expression of T7 polymerase for target protein production. This demonstrates that protein synthesis in the cells is directed at the XplA flavodoxin. The synthesised XplA flavodoxin may be incorrectly folded and susceptible to proteolysis, explaining its absence from samples analysed using SDS-PAGE. However, whether or not the *xplA-FMN* is expressed remains unclear at this point.
Figure 5.18: Time-lapse production of XplA flavodoxin. (A) 15\% (w/v) SDS-PAGE gel showing the expression profile of XplA flavodoxin in HMS174 (DE3) cells carrying the pET15b(xplA-FMN) construct. The gel shows (1 mL) samples taken at several time points from cells induced with 100 µM IPTG, and from cells not induced with IPTG as the control. The gel shows from left to right; molecular weight marker (1), time = 0 (2), 1 h control (3), 1 h induced (4), 2 h control (5), 2 hour induced (6), 4 h control (7), 4 h induced (8), 24 h control (9) and 24 h induced (10). (B) 15\% (w/v) SDS-PAGE gel showing the expression profile of XplA flavodoxin in HMS174 (DE3) cells carrying the pET11a(xplA-FMN) construct. Gene expression was induced as with the pET15b(xplA-FMN) construct, and the gel layout is as with gel (A).

Although xplA-FMN expression was not observed during small scale expression trials, expression was scaled up to 0.5 L cultures. The soluble cell lysate was purified using Ni-NTA affinity chromatography, and the fractions analysed by SDS-PAGE gel (Figure 5.19). However, despite the appearance of bands of the approximate size of XplA flavodoxin (16.6 kDa including N-terminal His\textsubscript{6}-tag), subsequent identification of those bands by mass spectrometry were all negative with respect to identifying XplA.
flavodoxin. Furthermore, the protein was not found in the insoluble fraction. Scaled up expression of xplA-FMN using the pET11a vector and purification using DEAE ion exchange chromatography also showed that the protein was absent from both the soluble and insoluble fractions.

Figure 5.19: Ni-NTA affinity chromatography purification of XplA flavodoxin. 15% (w/v) SDS-PAGE gel showing XplA flavodoxin purification using Ni-NTA affinity chromatography. The gel shows from left to right; insoluble fraction (1) molecular weight marker (2), soluble fraction of cell lysate (3), eluate from a 10 mM imidazole wash (4) and 1 mL fractions from a 300 mM imidazole wash (5-9). Expression conditions were: 100 µM IPTG, 20 °C, 24 h using pET15b(xplA-FMN) in HMS174 (DE3) cells.

5.9: Summary

Following extensive expression trials using a variety of E. coli expression strains and vectors carrying different fusion tags, soluble XplB was overexpressed with an N-terminal His6-tag using pET15b and also as an fusion protein with an N-terminal TF protein using pCold TF. In both constructs the protein was isolated with minimal FAD incorporated. Clearly the His6-tagged XplB is produced with the FAD cofactor incorporated. However, separation of the recombinant protein from the cell lysate facilitates the dissociation and separation of the FAD cofactor when the protein is loaded onto a chromatography column. In contrast, the FAD does not appear to be incorporated during protein synthesis of the TF-XplB fusion, as in addition to the absence of FAD in the purified fusion protein, the FAD was not observed in the cell lysate or the buffers from chromatographic steps, which would have been unusually yellow in the presence of free FAD. The production of the relatively large TF protein
(48 kDa) prior to the synthesis of the attached XplB could hinder FAD incorporation, if FAD binding is subsequent to protein folding. Experiments to identify buffer conditions that encouraged the retention FAD and all other reconstitution efforts were unsuccessful, and the holoprotein is yet to be obtained.

Biophysical characterisation of the XplB apoprotein using CD spectroscopy revealed the presence of predominately β-sheet secondary structural content, lacking significant α-helix features. The FUV CD spectrum of XplB did not share similarities with previously reported CD spectra for flavoproteins [363]. This supports the earlier suggestion that the stability of the protein is compromised following purification, and that this is the reason for the FAD loss observed, rather than weak FAD affinity in the native enzyme. MALLS data showed that XplB was monomeric and constituted a single population of proteins in solution. However, the molecular weight derived from MALLS data was low compared with that observed on a SDS-PAGE gel and the predicted molecular weight from the amino acid sequence. A possible explanation is that the molecular weight derived from MALLS data is inferred from the $R_h$, and the fact that XplB may not be correctly folded leads to this discrepancy.

Protein and FAD interactions were not detected using spectroscopic and fluorescence techniques, or by ITC. However, XplB showed NADPH oxidation activity in the presence of free FAD, indicating that binding does occur, and when bound to protein the FAD is reduced to the HQ state (as shown by UV/Visible spectroscopy). Furthermore, XplB was shown to be able to reduce its partner enzyme XplA in the presence of free FAD and when supplied with electrons from NADPH to support RDX reduction. Multiple amino acid sequence alignment with some of the most related microbial oxidoreductases showed that XplB shared some sequence homology with AdR-like enzymes. XplB retained most of the residues identified in the bovine AdR structure that are important for FAD binding. However, due to the diversity of these oxidoreductases, the identification of other residues that are important for FAD binding and whether they were retained in XplB was not clear.

Flavin loss was also observed from XplA, which was clearly visible during purification. The purified protein only contained ~one third of the content of FMN for molecule of heme, suggesting weak FMN affinity. Interestingly, *E. coli* MioC, which is closely related to the XplA flavodoxin, when overexpressed and purified from its
Chapter 5: Results III

A host bacterium was reported as being only ~25% holoprotein, suggesting relatively weak FMN affinity in this protein [382]. However, in contrast to XplB, reconstitution of XplA with FMN was achieved by altering the buffers to incorporate moderate concentrations of NaCl (~150 mM). Subsequently, using fluorescence spectroscopy, XplA was shown to bind FMN with a $K_d$ of 1.07 µM. This value is high compared to the FMN binding affinity of most bacterial flavodoxins, but comparable to the flavodoxin WrbA, which binds FMN with a $K_d$ of 2.07 µM and also purifies with incomplete FMN incorporation [356, 383]. Comparison of the amino acid sequence of XplA flavodoxin with those of similar microbial flavodoxins showed the retention of residues important for FMN binding, including the two tyrosines which form aromatic stacking interactions with the si- and re-sides, respectively, of the isoalloxazine ring.

Potentiometric analysis of the XplA showed that the midpoint reduction potentials of the FMN redox couples were very positive compared to other well characterised flavodoxins, with potentials of -80 and -170 mV for the OX/SQ and SQ/HQ couples, respectively. The high potentials of the FMN are consistent with the relatively high potential of the heme Fe$^{3+}$/Fe$^{2+}$ and thus may be important to facilitate rapid electron transfer through the enzyme, and also possibly reflect the proposed evolution of the XplA enzyme as a reductase. Furthermore, the very high potential of the OX/SQ couple suggests that it is unlikely to be the catalytically relevant species. It is reasonable to assume that substrate binding increases the heme Fe$^{3+}$/Fe$^{2+}$ reduction potential by ~100 mV, which would elevate the substrate-bound heme Fe$^{3+}$/Fe$^{2+}$ potential to ~150 mV, and thermodynamically well positioned to receive electrons from the SQ/HQ couple.

The XplA flavodoxin domain could not be expressed to produce measurable amounts of protein on a SDS-PAGE gel, and no XplA flavodoxin was purified independently of the P450 partner. In light of the issues with weak FMN affinity of the intact XplA, we believe that the XplA flavodoxin is not particularly stable when expressed as a recombinant protein, and subsequent expression of the XplA flavodoxin independent of the P450 leads to the production of a partially folded polypeptide which is susceptible to proteolysis by proteases. However, with more extensive expression trials of the XplA flavodoxin to incorporate larger fusion tags, including GST or thioredoxin, this flavodoxin may yet be isolated independent of the P450 domain, as it is clearly functional when purified as part of the intact XplA. Given that it is stable...
when expressed fused to the P450 portion of XplA in its “native” form, we believe that expression of this flavodoxin will be possible as a fusion construct. Trials could also include new constructs of the XplA flavodoxin domain to incorporate a larger portion of the linker region separating the P450 and flavodoxin domains of XplA, i.e. residues 141-149.

Due to the issues with flavin loss observed in both flavoproteins of the XplA/B redox system, and the inability to express and purify the XplA flavodoxin independent of the P450, P450-redox partner and redox partner-redox partner interactions could not be analysed in detail in this novel redox system. Furthermore, for the same reasons, the biochemical and structural characterisations of these two flavoproteins remain incomplete. Many avenues of expressing soluble and fully functional XplB from *R. rhodochrous* have been attempted using this XplB species without successful isolation of the soluble holoprotein. Interestingly, an XplB homologue is present in *Gordonia* (sp. KTR9), as part of a gene operon also consisting of an *xplA* gene and CYP151C. This XplB is a natural fusion enzyme, which is fused to ~70-80% of a GlnS-like enzyme, which might make it more amenable to heterologous protein production in *E. coli* systems [299]. Given the inability to purify soluble XplB from *R. rhodochrous* with FAD cofactor bound and due to the high degree of similarity with XplB from *Gordonia* (99%), the latter XplB may represent a more appealing and tractable protein to synthesise, purify and characterise. Successful expression of this XplB form in a soluble and functional state would then facilitate the introduction of cleavable sites in the linker region between XplB and GlnS, which would in turn allow isolation of XplB as an independent protein for further structural and biochemical analysis.

Alternatively, the *R. rhodochrous xplA-xplB* genes could be engineered to express both as an XplA/B fusion enzyme, with XplB fused to the N-terminus of XplA to maintain the natural flow of electrons as the FMN-binding domain is located to the N-terminal of the P450. These fusions are not unprecedented, as rat CPR has been successfully fused to several mammalian P450s including bovine adrenal CYP17A1, and microsomal CYP4A1 (rat) and CYP3A4 (human) to create enzymatically active fusion enzymes [384, 385]. The XplA-HD has also been fused to a variety of reductases and displayed catalytic activity [386]. An XplA/XplB fusion may stabilise XplB in the FAD-bound conformation. However, there is the possibility of the fusion protein retarding
Chapter 5: Results III

catalytic efficiency due to the locking of the proteins in conformations preventing relevant interactions between XplA and XplB, and/or between XplB and pyridine nucleotide coenzymes.
Chapter 6: Conclusions and Future Perspectives
Chapter 6: Conclusions

6.1: Introduction

The work presented in this thesis concerns the characterisation of two enzymes from a plasmid encoded gene operon in the soil bacterium R. rhodochrous, which imparts the ability in this bacterium (and in several related and unrelated bacteria also possessing the genes) to degrade the environmental pollutant and military explosive compound, RDX. One of the enzymes, called XplA, belongs to a well characterised superfamily of heme-\(b\) containing enzymes called the cytochromes P450. The other, its cognate flavodoxin reductase partner named XplB, is an enzyme belonging to the FAD-binding oxidoreductase family of enzymes. The XplA enzyme of this pair is unique amongst the P450 family owing to its domain organisation, which in addition to the heme domain also contains a fused flavodoxin protein, containing a non-covalently bound FMN cofactor. Furthermore, the degradation pathway of RDX breakdown has been shown to proceed via initial reductive denitration steps catalysed by XplA, which although uncommon for a P450 enzyme, is not unprecedented, e.g. CYP3A from rat liver microsomes is able to reduce glyceryl trinitrate to form nitric oxide [387]. Glyceryl trinitrate is also an explosive but has medical uses, and it is similar to RDX in that it has three exposed nitro groups. However, given the generalist nature of microsomal P450s, which are known to display ligand promiscuity, the fact that CYP3A bound glyceryl trinitrate containing nitro groups that are prone to reduction as a substrate, and catalysed its reduction was unsurprising. Therefore, it is interesting that a bacterial P450 appears to have evolved under the evolutionary influence of RDX to adapt as a reductase. Therefore, the XplA enzyme can be deemed interesting due to its construction, and also its functionality.

6.1.1: Spectroscopic Characterisation of XplA

The XplA enzyme was successfully isolated in a fully functional state. However, initial spectroscopic analysis of the purified protein revealed FMN loss, and imidazole contamination resulting in heme iron coordination by the ligand. This was later shown to be due to XplA’s unusually weak affinity for FMN, and its extraordinarily strong affinity for imidazole. Comparison of the amino acid sequence of the XplA flavodoxin domain with some of the most related microbial flavodoxins failed to
Chapter 6: Conclusions

identify immediate differences in conserved residues which may contribute to the weak FMN binding observed in XplA. However, buffer conditions that encouraged FMN retention were later identified, and the use of this buffer in all subsequent experiments ensured that the enzyme retained >90% FMN incorporation on purification. The problem of imidazole contamination was remedied by extensive dialysis of Ni-NTA affinity chromatography purified protein, and later by purifying XplA using DEAE ion exchange and hydroxyapatite chromatography. The use of the latter method eliminated the need to reconstitute the protein with FMN and for dialysis to remove imidazole, and allowed the isolation of pure and exogenous ligand-free XplA holoprotein.

The purified enzyme displays a UV/Visible spectrum characteristic of most P450s, and, as expected for a P450-flavodoxin fusion enzyme, having absorbance features at ~360 nm and between 450-500 nm corresponding to the wavelengths of maximum absorption of the oxidised FMN cofactor. In contrast to previously published spectroscopic data on RDX substrate binding, which was non-standard for a LS to HS heme transition normally associated with substrate binding to a P450 enzyme, our titrations produced spectra with induced spectral changes consistent with the spectral changes normally observed for substrate binding to a P450 enzyme, with the Soret peak shifting from 417 nm, corresponding to the LS ferric heme, to 396 nm corresponding to the HS ferric heme as the substrate concentration was increased. Furthermore, binding curves derived from the titrations yielded a $K_d$ of 7.47 μM for RDX binding, which is significantly tighter than the previously reported binding affinity. The most likely explanation for this, and given the unusual absorption spectrum of the XplA sample from which the binding affinity was originally derived [106], is that apparent weak affinity is due to insufficient removal of imidazole prior to binding studies with RDX. Strong evidence for the presence of imidazole is observed in the published spectrum, due to the unusual spectral changes described above and also due to the position of the Soret peak prior to addition of substrate, which was at 421 nm. This is unusually long for a P450 enzyme in the resting state with water ligated as the 6th axial ligand. The higher than expected absorption near ~365 nm is also consistent with a heme species partially ligated to a nitrogenous ligand, possibly imidazole or even Tris. In addition, XplA’s affinity for imidazole is so strong that binding was also observed for the ferrous heme. Most P450s bind imidazole with $K_d$ values in the high mM range, and no $K_d$ value for imidazole binding
to a P450 in the ferrous state has previously been reported. The UV/Visible spectrum of this ferrous XplA-imidazole complex displays novel spectral features with Soret maximum at 439 nm and weaker absorbance peaks at 538 and 566 nm, respectively.

Other non-standard spectroscopic observations were also made for XplA, including binding to ligands atypical to P450s, such as Tris and a variety of sulphur-containing ligands. The active sites of P450 enzymes are mostly hydrophobic, therefore binding of polar molecules is usually disfavoured, but the active site structure in XplA is consistent with the unusually strong affinity for imidazole demonstrated here for XplA. The UV/Visible spectra of the XplA-DTT and XplA-BME complexes are similar to those previously reported for P450cam bound to p-chlorothiophenol, the only other P450 for which binding of sulphur-containing ligands has been detailed, with an unusual split Soret (hyperporphyrin) with peaks at ~374 and 453.5 nm corresponding to ligation of the sulphur ligands in the thiolate state [305]. A third peak also arises at 423.5 nm due to ligation of the sulphur ligand in the thiol state, and the relative intensities of this peak and the split Soret peaks can be manipulated by altering the pH, consistent with reversible protonation/deprotonation of the proximal ligand. These unusual spectroscopic and ligand binding properties were confirmed with EPR and MCD spectroscopy, which provided unique spectroscopic fingerprints for these enzyme-ligand complexes. In addition, the tight binding of RDX at the heme distal face and the reductive reaction catalysed by XplA is highlighted by the ability of RDX to oxidise the Fe(II)CO complex of XplA under anaerobic conditions, presumably due to RDX displacement of the CO ligand and efficient electron transfer from the ferrous iron to denitrate RDX. CO is often deemed as an irreversible heme inhibitor, therefore the observation that a substrate could displace it was unexpected. The CO binding affinity was determined using stopped-flow techniques, and shown to be weaker compared to the binding affinities for CO demonstrated by other P450s due to a high dissociation rate. This likely contributes to the observation of RDX oxidation of CO-bound ferrous heme.

The same techniques were also used to characterise the oxyferrous complex of XplA which shared spectral features to the oxyferrous complexes previously reported for other P450s [318, 388, 389], although the immediate ferric-peroxo and ferric hydroperoxide intermediates following the second electron reduction and protonation steps could not be observed using the full length enzyme. The prospect
of these intermediates being observed was rationalised due to the presence of the flavodoxin domain, pre-reduced by sodium dithionite, which could rapidly transfer the second electron following dioxygen binding after rapid mixing of reduced enzyme with dioxygen. The characterisation of the oxyferrous complex might suggest XplA retains capacity to perform oxidative reactions on other substrates. The production of indigo in *E. coli* cells expressing *xplA* hints at this capacity, although indole binding to XplA was not detected spectrophotometrically, and its transformation to indigo by XplA was not seen *in vitro*.

Analysis of the thermodynamic properties of the enzyme using potentiometric titrations with sodium dithionite/potassium ferricyanide also revealed relatively high midpoint reduction potentials for the heme Fe3+/Fe2+, FMN OX/SQ and SQ/HQ redox couples with potentials of -267, -80 and -172 mV (versus NHE), respectively. Previous studies of Phe393 mutants in the P450 BM3 enzyme indicated that positive shifts in the heme iron potential were correlated with faster heme reduction by the redox partner, but also with greater oxyferrous complex stability, diminished driving force for electron transfer to bound dioxygen from ferrous heme iron, and much less efficient oxidative catalysis [14, 390]. Thus, the positive potential of the XplA heme iron may be an important clue regarding its evolution into a predominantly reductive catalyst. The FMN (SQ/HQ) couple is the likely electron donor to the XplA heme, and natural mutations to the FMN binding site that have resulted in the relatively positive potential of this couple may have enabled efficient electron transport between XplB, the XplA FMN and the heme, while simultaneously resulting in diminished FMN affinity for the protein.

### 6.1.2: Biophysical and Structural Characterisation of XplA

Biophysical characterisation of XplA, and the separated XplA-HD free from the flavodoxin showed that both proteins were could be purified to homogeneity and were stable. Importantly they were both monodisperse, demonstrating their suitability for crystallographic studies. Crystals were obtained from crystallisation experiments with the intact protein. However, these crystals were very unstable and deteriorated before they could be frozen. The conditions which produced the crystals were not reproducible. Presumably success in obtaining crystals of “intact” XplA
were due either to the superior quality of that sample of intact protein, or because sufficient amounts of protein were cleanly proteolysed to promote crystal growth of a cleaved construct.

The structure determined from co-crystallisation of the XplA-HD with RDX showed that imidazole was bound to the heme, demonstrating that complete removal of imidazole was not achieved, consistent with the previously reported structure, which was also in an imidazole-bound form. Close inspection of the distal pocket enabled rationalisation of the reason for the extraordinarily high affinity for imidazole. In addition to the Fe-imidazole nitrogen bond, an extended hydrogen bonding network to the other imidazole nitrogen and steric complementarity of residues around the imidazole enable its near-linear geometry of coordination to the iron. The space for imidazole to bind in this geometry is partly due to the severe kink in the I helix which removes steric constraints which otherwise would have forced the imidazole to bind in a mode more commonly observed for most P450s. The structural features introduced by this unexpected conformational change and other residues around the heme active site create a binding pocket which makes favourable interactions with imidazole on all sides, and to a lesser extent to the other small polar molecules shown to bind XplA in this project. In addition to the bound imidazole, a molecule with electron density that could be PEG is also present in the heme active site channel, forming a ternary complex. The PEG makes van der Waals contacts with the imidazole, and with residues lining the hydrophobic active site. The binding of a PEG molecule, which shares similar structure to some fatty acids, provides clues to an evolutionary origin for XplA, possibly pointing to a fatty acid binding role for a progenitor of this RDX reducing enzyme.

Further optimisation of the purification protocols eventually lead to the isolation of XplA-HD without the use imidazole, and crystals were also obtained from protein purified using this method. Although the crystals appeared to contain the enzyme in the substrate-bound HS form, they did not diffract to the resolution necessary to determine a structure. The data did, however, show the presence of a bound ligand, and the electron density of this ligand was not consistent with that expected for water, RDX or imidazole.

As the structure of the full length enzyme remains elusive, and the structure of the flavodoxin domain is also uncharacterised, it remains unclear as to why XplA has
weak affinity for FMN. Furthermore, attempts to express the flavodoxin domain independently were unsuccessful, pointing to a lack of stability of this module in absence of its fused P450 partner.

6.1.3: Characterisation of XplB

Following extensive expression and purification trials of XplB in various constructs, soluble and functional protein was recovered using a His$_6$-tagged construct. However, the purified protein was only functional when supplemented with FAD, as XplB experienced extensive FAD loss during the purification steps. This indicates the protein is synthesised with FAD incorporated, and that the stability of the protein is compromised during purification leading to FAD loss. Biophysical characterisation of the purified protein using CD spectroscopy confirmed the protein may be incorrectly folded, as the CD spectra produced by XplB were not consistent with those previously reported for correctly folded flavoproteins. In particular, the FUV CD spectrum shows the purified XplB to consist almost completely of β-sheets and lacking significant α-helical content. This contrasted with the structures of proteins belonging to the oxidoreductase family previously characterised, which all share a variation of the Rossmann fold containing significant α-helical and β-sheet secondary structural content.

Although analysis of the purified protein using SDS-PAGE showed the protein had been partially clipped leading to the presence of a major and a smaller band. MALLS showed that XplB existed as a homogenous monodisperse species in solution. There was a small variation in the molecular masses derived from the MALLS data, likely due to the presence of the clipped and unclipped versions. Regardless of the folded state of the purified protein, it was shown that XplB could interact with NADPH to support the activity of XplA in driving catalytic turnover of RDX, although the rate observed is likely to be much lower than the true catalytic rate, due to non-productive interactions of FAD-free XplB with XplA and the equilibrium between FAD-free and FAD-bound forms of XplB. The productive interactions showed that although binding of the oxidised FAD cofactor to XplB was very weak and undetectable using the methods tested including fluorescence spectroscopy and ITC, trapping of the FAD HQ demonstrated that the purified XplB does bind the FAD cofactor (Chapter 5; 5.5.5.1
Chapter 6: Conclusions

Figure 5.10). Although the activity of XplB was demonstrated, the inability to isolate the XplB holoprotein with FAD fully incorporated has restricted further biochemical characterisation of this flavodoxin reductase, such as kinetic studies of the XplA-XplB interactions and their electron transfer steps. Such kinetic analyses would be compromised by the heterogeneity of the XplB samples.

6.1.4: Future Work

Further work on the XplA enzyme should include efforts to produce a structure of the isolated heme domain with ligands bound other than imidazole. Substrate- and product-bound heme domain structures may show mechanistically revealing details of RDX breakdown. Preceding this, it would first be necessary to confirm binding of the product molecules NDAB and MEDINA by performing optical titrations with XplA. Structural studies would be best attempted by optimising conditions under which the alternative crystal form is obtained, using protein purified by DEAE ion exchange and hydroxyapatite chromatography, and which could be used successfully for X-ray diffraction experiments. To aid these efforts, crystallisation of the ligand-free protein should be attempted, which would enable subsequent soaking with ligands of interest.

With regards to further studies on the flavodoxin domain of XplA, the first step should be to develop a construct to express the protein fused to a larger tag than has been attempted to date with a His6-tag, to stabilise and protect the recombinantly fused protein from proteolysis, as appears to occur during expression trials, through effective single step purification. These tags may include GST- or thioredoxin. Constructs of the XplA flavodoxin should also be generated to incorporate a larger section of the linker region prior to the start of the P450 domain. As for the XplB from R. rhodochrous, it appears that several avenues to express this protein have been explored, with only limited progress made in the isolation of this highly insoluble protein. As a homologue has been identified in Gordonia (sp. KTR9), which is fused to a large portion of a GlnS enzyme, efforts should be made to express and purify this homologue which may be more amenable to heterologous recombinant protein expression using E. coli to produce soluble XplB. This XplB is part of a larger operon consisting of two P450 enzymes, one of which is also XplA. The isolation of XplB
would allow further characterisation of the kinetics of the interactions between these two enzymes, and allow the determination of the individual rate constants for e.g. the NAD(P)H-dependent reduction of the XplB FAD, and for electron transfer between reduced FAD the XplA FMN and heme cofactors. These experiments would involve pre-steady state kinetic experiments using stopped-flow absorption/fluorescence techniques. The isolation of the XplA flavodoxin independent of the heme domain would also facilitate such studies, and complete the characterisation of both enzymes from the XplA/XplB redox system.

6.1.5: Final Remarks

In conclusion, this project builds on the preliminary biochemical and structural characterisation of XplA previously reported. The goal of the project was to express and purify both components of this novel P450 redox system. Successful isolation of both enzymes would facilitate the analysis of the interactions between the enzymes and advance the pursuit of other major goals of the project – specifically to structurally characterise the full length XplA and a substrate-bound version of the enzyme, and to interrogate previously proposed mechanisms of P450-redox partner interactions and RDX breakdown pathways, respectively. Despite problems associated with characterisation of XplB and in obtaining structural detail for intact XplA, other interesting properties of this novel enzyme system were elucidated, including the determination of XplA’s high affinity for imidazole, which could be rationalised through analysis of the active site of the imidazole-bound structure of the heme domain of XplA, and also of the XplA flavodoxin domains weak affinity for FMN, which initially necessitated its reconstitution with the FMN cofactor following purification. Conditions were later established which encouraged better FMN retention, but the reasons for the weak FMN affinity for XplA remain unclear. However, structural changes near the FMN-binding environment leading to the elevated redox potentials determined of the FMN cofactor may have also weakened XplA’s affinity for FMN. Subsequently, a relatively positive redox potential for the XplA heme was also determined. These and other modifications in XplA including the absence of the conserved acid/alcohol pair almost absolutely conserved in most other P450s likely reflect XplA’s specialisation as a reductase, although clearly the enzyme retains capacity to also function as an oxygenase. In addition, the characterisation of
Chapter 6: Conclusions

the XplB flavodoxin reductase was advanced by recovery of this previously intractable protein in a soluble and homogenous state, which was able to interact with its redox partners. However, this enzyme also demonstrated very weak FAD affinity, and the reasons for this remain unclear. The properties of this enzyme and also of XplA, including the capacity of the latter in functioning as oxygenase and other interesting properties remain to be explored in this novel enzyme system. Methods to study some of these properties have been proposed in the preceding section.
References


References


References

References


References


References


References


References


References


263. Stratagene, QuikChange® II Site-Directed Mutagenesis Kit.


References


References

References


References


References


Addendum

**Figure A1: E. coli codon-optimised XplA DNA sequence**

```
CATATGACCGATGTAACCGTGCTTTCGGTACCGGAAACATGGTAGCTGACGACATCGCGTCGCTGCGCTGGGTGAATTCGATATCGAAGCGACGGTGGTTGGCATGGAGGACTTCGACGTTGCGGATCT
GGCAGC
```

**Figure A2: E. coli codon-optimised XplB DNA sequence.**

```
CATATGACCGATGTAACCGTGCTTTCGGTACCGGAAACATGGTAGCTGACGACATCGCGTCGCTGCGCTGGGTGAATTCGATATCGAAGCGACGGTGGTTGGCATGGAGGACTTCGACGTTGCGGATCT
GGCAGC
```
Table A1: UV/Visible absorption and binding affinity data for ligand binding to XplA. Spectral binding data and methods for determination of dissociation constants ($K_d$ values) for RDX substrate and inhibitors are given in Materials & Methods (2.2.14). “Fe$^{II}$” indicates that spectral details are for the reduced heme species. Unless indicated, spectra are for the ferric heme iron form of XplA. No spectral changes indicative of XplA binding were obtained on addition of econazole, fluconazole, ketoconazole, miconazole, clotrimazole, voriconazole, 1-phenylimidazole, histidine or TNB to XplA. However, the $K_d$ for TNB (198 ± 41 μM) was determined by competitive displacement of TNB using the RDX substrate. $K_d$ values for oxygen and carbon monoxide were determined using stopped-flow methods, as described in Materials & Methods (2.2.18). “sh” indicates an absorption shoulder rather than a discrete peak.

<table>
<thead>
<tr>
<th>XplA ligand bound</th>
<th>Soret maximum (nm)</th>
<th>Q-band maxima (nm)</th>
<th>$K_d$ value (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>417</td>
<td>540/566</td>
<td>-----</td>
</tr>
<tr>
<td>none (Fe$^{II}$)</td>
<td>408</td>
<td>542</td>
<td>-----</td>
</tr>
<tr>
<td>imidazole</td>
<td>424.5</td>
<td>543/575</td>
<td>1.6 ± 0.1</td>
</tr>
<tr>
<td>imidazole (Fe$^{II}$)</td>
<td>439</td>
<td>538/566</td>
<td>2730 ± 260</td>
</tr>
<tr>
<td>RDX</td>
<td>396</td>
<td>540 (sh)</td>
<td>7.5 ± 0.2</td>
</tr>
<tr>
<td>4-phenylimidazole</td>
<td>423</td>
<td>544.5/572 (sh)</td>
<td>0.25 ± 0.05</td>
</tr>
<tr>
<td>carbon monoxide (Fe$^{II}$)</td>
<td>446.5</td>
<td>548</td>
<td>2.6 ± 0.5</td>
</tr>
<tr>
<td>oxygen (Fe$^{II}$)</td>
<td>424.5</td>
<td>557</td>
<td>0.34 ± 0.30</td>
</tr>
<tr>
<td>dithiothreitol</td>
<td>374/423.5/453.5</td>
<td>550.5</td>
<td>70.5 ± 1.1</td>
</tr>
<tr>
<td>β-mercaptoethanol</td>
<td>372/423.5/453.5</td>
<td>552</td>
<td>58.5 ± 3.2</td>
</tr>
<tr>
<td>dimethyl sulphide</td>
<td>423</td>
<td>540/569 (sh)</td>
<td>81.4 ± 7.0</td>
</tr>
<tr>
<td>morpholine</td>
<td>422</td>
<td>540.5/569</td>
<td>31.8 ± 0.5</td>
</tr>
<tr>
<td>pyrrolidine</td>
<td>423.5</td>
<td>545/569 (sh)</td>
<td>481 ± 55</td>
</tr>
<tr>
<td>piperazine</td>
<td>422</td>
<td>541/570 (sh)</td>
<td>970 ± 40</td>
</tr>
<tr>
<td>cyanide</td>
<td>436</td>
<td>559</td>
<td>294 ± 20</td>
</tr>
<tr>
<td>nitric oxide</td>
<td>431</td>
<td>543/573</td>
<td>-----</td>
</tr>
<tr>
<td>Tris</td>
<td>423</td>
<td>541.5/574 (sh)</td>
<td>4570 ± 420</td>
</tr>
</tbody>
</table>
**Table A2: Melting temperatures of buffer screen of XplA-HD.** Buffers shown are from the JBS Solubility Kit part A. Buffers which did not produce a $T_m$ are not shown. Final concentrations for all buffers were 50 mM diluted from 100 mM stocks. The buffers are listed in descending order to the stability of XplA-HD.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH</th>
<th>$T_m$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium/potassium phosphate</td>
<td>7.0</td>
<td>44.7</td>
</tr>
<tr>
<td>Bis-Tris</td>
<td>6.0</td>
<td>41.3</td>
</tr>
<tr>
<td>Imidazole</td>
<td>8.0</td>
<td>39.2</td>
</tr>
<tr>
<td>Sodium citrate</td>
<td>5.5</td>
<td>37.9</td>
</tr>
<tr>
<td>Bis-Tris propane</td>
<td>6.5</td>
<td>37.4</td>
</tr>
<tr>
<td>Ammonium acetate</td>
<td>7.0</td>
<td>36.7</td>
</tr>
<tr>
<td>Tris</td>
<td>7.5</td>
<td>35.8</td>
</tr>
<tr>
<td>HEPES</td>
<td>7.5</td>
<td>35.6</td>
</tr>
<tr>
<td>Bicine</td>
<td>8.5</td>
<td>35.4</td>
</tr>
<tr>
<td>EPPS</td>
<td>8.0</td>
<td>35.4</td>
</tr>
<tr>
<td>MOPS</td>
<td>7.0</td>
<td>35.4</td>
</tr>
<tr>
<td>Sodium/potassium phosphate</td>
<td>6.0</td>
<td>35.4</td>
</tr>
<tr>
<td>Tris</td>
<td>8.5</td>
<td>35.4</td>
</tr>
<tr>
<td>CHES</td>
<td>9.0</td>
<td>35.2</td>
</tr>
<tr>
<td>CHES</td>
<td>9.5</td>
<td>34.8</td>
</tr>
<tr>
<td>ADA</td>
<td>6.5</td>
<td>32.3</td>
</tr>
<tr>
<td>MES</td>
<td>6.2</td>
<td>32.3</td>
</tr>
<tr>
<td>CAPS</td>
<td>10.0</td>
<td>32.1</td>
</tr>
<tr>
<td>Sodium acetate</td>
<td>4.5</td>
<td>31.2</td>
</tr>
<tr>
<td>PIPPS</td>
<td>3.7</td>
<td>31.1</td>
</tr>
<tr>
<td>Sodium/potassium phosphate</td>
<td>5.0</td>
<td>31.1</td>
</tr>
</tbody>
</table>

**Table A3: Melting temperatures of additive screen of native and ligand-bound XplA-HD.** Additives shown are from the JBS Solubility Kit part B. Additives which did not produce a $T_m$ are not shown. The $T_m$’s shown are for the ligand-free enzyme, and in complex with RDX and imidazole at 1 and 0.5 mM, respectively. The final concentrations of additives shown were diluted from 5x stocks. The additives are listed in descending order to the stability of the ligand-free enzyme.

<table>
<thead>
<tr>
<th>Additive</th>
<th>Concentration</th>
<th>$T_m$ (°C)</th>
<th>$T_m$ (°C) + RDX</th>
<th>$T_m$ (°C) + imidazole</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>80 mM</td>
<td>38.5</td>
<td>41</td>
<td>42.4</td>
</tr>
<tr>
<td>Glycerol</td>
<td>8%</td>
<td>37.3</td>
<td>39.3</td>
<td>41.7</td>
</tr>
<tr>
<td>NaCl</td>
<td>40 mM</td>
<td>37.4</td>
<td>39.3</td>
<td>41.8</td>
</tr>
<tr>
<td>BME</td>
<td>8 mM</td>
<td>37.2</td>
<td>38.7</td>
<td>40.5</td>
</tr>
<tr>
<td>DTT</td>
<td>0.8 mM</td>
<td>37.2</td>
<td>38.8</td>
<td>40.6</td>
</tr>
<tr>
<td>Glycerol</td>
<td>4%</td>
<td>37.2</td>
<td>39</td>
<td>40.8</td>
</tr>
<tr>
<td>DTT</td>
<td>4 mM</td>
<td>37.1</td>
<td>38.6</td>
<td>40</td>
</tr>
<tr>
<td>NaCl</td>
<td>16 mM</td>
<td>37.1</td>
<td>39</td>
<td>41</td>
</tr>
<tr>
<td>Control</td>
<td>N/A</td>
<td>35.9</td>
<td>38.6</td>
<td>39</td>
</tr>
<tr>
<td>Octyl glucopyranoside</td>
<td>0.08%</td>
<td>35.3</td>
<td>36.7</td>
<td>38.4</td>
</tr>
<tr>
<td>CHAPS</td>
<td>1.6 mM</td>
<td>35.2</td>
<td>35.6</td>
<td>37.2</td>
</tr>
<tr>
<td>Octyl glucopyranoside</td>
<td>0.80%</td>
<td>35.2</td>
<td>35.7</td>
<td>37.3</td>
</tr>
</tbody>
</table>