Characterisation of the unique *Campylobacter jejuni* cytochrome P450, CYP172A1.

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

2013

Peter C. Elliott
# Table of Contents

List of Figures ......................................................................................................................... 7  
List of Tables ............................................................................................................................ 10  
List of abbreviations ................................................................................................................ 11  
Abstract .................................................................................................................................. 14  
Declaration ................................................................................................................................. 15  
Copyright ................................................................................................................................... 15  
Acknowledgements ................................................................................................................... 16  
Chapter 1. Introduction ............................................................................................................. 17  
1.1.1. The *Campylobacter* genus and related organisms ...................................................... 17  
1.1.2. *Campylobacter jejuni* and its general features .............................................................. 18  
1.1.3. *C. jejuni* infection sources ........................................................................................ 20  
1.2. *C. jejuni* infection and symptoms .................................................................................. 21  
1.2.1. *C. jejuni* adhesion to epithelial cells ........................................................................... 21  
1.2.3. *C. jejuni* invasion of epithelial cells ........................................................................... 22  
1.3. Antibiotic resistance in *Campylobacter* species ............................................................... 23  
1.4. Surface structures of *C. jejuni* ...................................................................................... 25  
1.4.1. The Flagellum ............................................................................................................... 25  
1.4.2. *C. jejuni* lipooligosaccharide (LOS) ......................................................................... 28  
1.4.3. *C. jejuni* capsule polysaccharide (CPS) locus ............................................................ 30  
1.4.4. *C. jejuni* CPS structural variability and serotyping .................................................... 30  
1.5. The presence of CYP172A1 in *Campylobacter jejuni* ...................................................... 34  
1.6. History of the cytochrome P450 ...................................................................................... 36  
1.6.1. The prevalence of P450s in nature .............................................................................. 38  
1.6.2. Bacterial P450s ............................................................................................................ 40  
1.7. P450 Structure ................................................................................................................. 40  
1.8. Catalytic cycle of P450s ................................................................................................... 42  
1.9. Redox partners ................................................................................................................ 48  
1.10. Flavin-containing proteins ............................................................................................. 50  
1.10.1. Ferredoxin/Flavodoxin reductases ............................................................................. 52
Chapter 2. Materials and Methods ........................................................................58
2.1. Materials ......................................................................................................58
2.2. Growth media ...............................................................................................58
2.3. Antibiotics ....................................................................................................61
2.4. Campylobacter jejuni strains .......................................................................62
2.5. Purification of C. jejuni genomic DNA .......................................................63
2.6. Amplification of target DNA by PCR .........................................................63
2.7. Preparation of electrocompetent C. jejuni ..................................................64
2.8. Sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis ... 65
2.9. Western blotting ...........................................................................................65
2.10. C. jejuni growth assays ............................................................................67
2.11. C. jejuni autoagglutination assays ..........................................................67
2.12. Electron Microscopy (EM) ........................................................................68
2.13. C. jejuni motility Assay ............................................................................68
2.14. Minimum Inhibitory Concentration (MIC) assays for C. jejuni ...............68
2.15. Polymyxin B susceptibility assays ..............................................................69
2.16. Disc diffusion assay for C. jejuni .............................................................70
2.17. Tricine-SDS-PAGE for C. jejuni capsule and LOS visualisation .............70
2.18.1. Fourier Transform Infrared (FT-IR) Spectroscopy ................................71
2.18.2. Cluster Analysis of FT-IR ..................................................................72
2.19. Surface Enhanced Raman Spectroscopy (SERS) .......................................73
2.20. Bacterial cell strains .................................................................................74
2.21. Restriction enzyme digests .......................................................................74
2.22. E. coli transformations ..............................................................................75
2.24. Site Directed Mutagenesis (SDM) ............................................................76
2.25. Preparation of E. coli glycerol stocks .........................................................77
2.26. E. coli expression trials of the cj1411c gene product ................................77
2.27. CYP172A1 solubility trials ........................................................................78
2.28. Temperature trials for CYP172A1 expression in E. coli ............................................ 78
2.28.1 IPTG gene induction trials ......................................................................................... 79
2.28.2. Growth Medium trials for target P450 production ................................................. 79
2.29. Expression and purification of CYP172A1 .................................................................. 79
2.30. CYP172A1 antibody production................................................................................. 80
2.31. UV-Visible spectroscopy ............................................................................................. 81
2.31.1. Dual beam UV-Visible spectroscopy ....................................................................... 81
2.31.2. Fluorescence spectroscopy ....................................................................................... 81
2.32. Pyridine Hemochromogen Assay .............................................................................. 82
2.33. Redox potentiometry ................................................................................................. 82
2.34. Differential Scanning Calorimetry (DSC) ................................................................. 84
2.35. Electron Paramagnetic Resonance (EPR) Spectroscopy ........................................ 84
2.36. Circular Dichroism (CD) spectroscopy ...................................................................... 85
2.37. Multiangle Laser Light Scattering analysis ............................................................... 85
2.38. Thermofluor assay for P450 protein stability analysis ........................................... 86
2.39. CYP172A1 crystallisation trials ................................................................................... 88
2.40. Protein structure modelling ......................................................................................... 89
2.41. Ligand binding assays ................................................................................................. 89
2.42. Compound Screening ................................................................................................. 90
2.43. P450 CO trapping experiments .................................................................................. 90
2.44. Enzyme turnover studies ............................................................................................ 91
2.45. Ion trap liquid chromatography (LC) ........................................................................ 92

Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni* 93

3.1. Introduction .................................................................................................................. 93
3.2. Construction of *cj1411c* insertion knockouts (PE2 and PE5) in *C. jejuni* 11168H and 81-176 strains ........................................................................................................ 94
3.3. Construction of the genetically complemented strains (PE3 and PE6) in *C. jejuni* 11168H and 81-176 .............................................................................................. 98
3.4. Detection of CYP172A1 in *C. jejuni* by probing western blots with CYP172A1 specific antiserum ............................................................................................................... 100
3.5. Growth comparisons of *C. jejuni* strains PE1-PE6 under microaerobic conditions at 42 °C ......................................................................................................................... 102
3.6. Investigating the effect of CYP172A1 on autoagglutination (AAG) ......................... 104
3.7. The role of CYP172A1 in the morphology of *C. jejuni* examined using electron microscopy................................................................. 107
3.8. The role of CYP172A1 in flagella-mediated motility ........................................ 110
3.9. The role of CYP172A1 in determining the sensitivity of *C. jejuni* to antimicrobial agents .......................................................................................... 112
3.10.1. The role of CYP172A1 in sensitivity to polymyxin B, using E-Test strips......... 114
3.10.2. The role of CYP172A1 in sensitivity to polymyxin B, using agar dilution........ 114
3.11. The role of CYP172A1 in the sensitivity of *C. jejuni* to antifungal azole compounds. 117
3.12. The role of CYP172A1 in LOS biosynthesis in *C. jejuni* ..................................... 119
3.13. The role of CYP172A1 in the biosynthesis of CPS in *C. jejuni* .......................... 121
3.14. Fourier Transform Infrared (FT-IR) spectroscopy of *Campylobacter* species ...... 124
3.15. Discrimination of *C. jejuni* PE1-PE6 strains by FT-IR spectroscopy............... 127
3.16. Probing the surface of *C. jejuni* using Surface Enhanced Raman Spectroscopy (SERS) ................................................................. 130
3.17. Summary ........................................................................................................ 134
Chapter 4. Expression, purification and biophysical characterisation of CYP172A1 .......... 140
4.1. Introduction ........................................................................................................ 140
4.2. Construction of the CYP172A1 recombinant protein expression system.......... 141
4.3. Expression trials of CYP172A1 in multiple *E. coli* strains ................................. 143
4.3.1. Investigating the effects of varying the temperature at which CYP172A1 expression is induced ................................................................. 144
4.3.2. Analysis of the effects of varying IPTG levels and the growth medium used to express CYP172A1 ........................................................................ 145
4.4. Large scale expression and purification of CYP172A1 ........................................ 148
4.5. Initial spectroscopic characterisation of CYP172A1 ........................................... 151
4.6. Calculation of the molar extinction coefficient for CYP172A1 using the pyridine hemochromagen determination method ................................................. 152
4.7. Determination of the heme coordination state in CYP172A1 by Electron Paramagnetic Resonance (EPR) spectroscopy .............................................. 155
4.8. Analysis of the tertiary structure of CYP172A1 using fluorescence spectroscopy .... 155
4.9. Analysis of the secondary structures of CYP172A1 .......................................... 159
4.10. Measuring CYP172A1 thermostability using Differential Scanning Calorimetry (DSC) .................................................................................................. 160
4.11. Analysis of the molecular weight and polydispersity of CYP172A1 by MALLS ...... 160
4.12. Use of the Thermofluor probe to identify stabilising buffers for CYP172A1 ..........163
4.12.1. Analysis of the effect of ligand binding on the thermostability of CYP172A1 ..........167
4.13. Investigations into the causal effects of CYP172A1 self aggregation ..................169
4.13. Effects of reducing agents on CYP172A1 aggregation ..................................169
4.14. Cleavage of the N-terminal His6 purification tag from CYP172A1 ....................172
4.15. Mutational truncation of full length CYP172A1 ........................................174
4.15.1. Expression trials of truncated versions of CYP172A1 ................................178
4.16. Structure determination of CYP172A1 .........................................................182
4.17. Summary .......................................................................................................185

Chapter 5. Ligand binding to CYP172A1 ..................................................................189
5.1. Introduction ......................................................................................................189
5.2 Type II binding molecules ...............................................................................190
5.2.1 Azole binding to P450s ................................................................................191
5.2.2. Further type II binding ................................................................................198
5.2.2.1 Binding of phenyl imidazoles (PIMs) to CYP172A1 .................................202
5.2.2.2 Binding of fatty acid linked imidazoles to CYP172A1 ...............................205
5.3. Type I ligands ...................................................................................................206
5.4 Compound screening to identify new ligands for CYP172A1 ............................208
5.4.1 Type I hits for CYP172A1 .............................................................................209
5.4.2 Type II hits for CYP172A1 ............................................................................209
5.4.3. Analysis of the interaction of novel type I and II (ChemDiv) compounds with CYP172A1 ..........................................................211
5.5. EPR of ligand-bound forms of CYP172A1 ......................................................220
5.6. Redox potentiometry of CYP172A1 ..................................................................224
5.7. CO trapping of ferrous CYP172A1 in the presence of redox partners ...............228
5.8. Turnover of ChemDiv compounds by CYP172A1 ...........................................237
5.8.1. Turnover of 204073 .....................................................................................238
5.8.2 Turnover of 213071 .....................................................................................238
5.9. C. jejuni sensitivity to 213071, 204074 and 204591 compounds .......................240
5.10. Summary ........................................................................................................242

Chapter 6. Conclusions .............................................................................................247
6.1. Introduction .......................................................................................................247
6.2. Initial investigations into the role of CYP172A1 in *C. jejuni* ........................................... 247
6.3. Expression, purification and biophysical characterisation of CYP172A1 .................................. 250
6.4. Ligand binding to CYP172A1 ................................................................................................ 252
6.5. Future directions ..................................................................................................................... 254
6.6. Concluding remarks ............................................................................................................. 255

References ........................................................................................................................................ 257
Appendix .......................................................................................................................................... 276

Word count 51,174

List of Figures

Figure 1.1. Rectangular cladogram of the selected members of the Campylobacterales order. 18
Figure 1.2. Scanning electron micrograph of *C. jejuni*. 19
Figure 1.3. Transmission routes of *C. jejuni* between hosts. 20
Figure 1.4. Model of *C. jejuni* internalisation via the activation of the Rho GTPase Cdc42. 23
Figure 1.5. Diagrammatic representation of the outer surface structures of *C. jejuni*. 26
Figure 1.6. Representation of the current model of the Pgl pathway in *C. jejuni*. 28
Figure 1.7. Sugar composition of the complete LOS core of *C. jejuni* 11168. 29
Figure 1.8. Location of *cj1411c* in the *C. jejuni* genome. 34
Figure 1.9. Diagrammatic representation of the *C. jejuni* P450 gene (*cj1411c*) and the surrounding genes aligned with orthologous genes from related bacteria. 35
Figure 1.10. Diagrammatic representation of the overall fold of P450\textsubscript{cam} (CYP101A1). 41
Figure 1.11. Heme binding motifs of selected bacterial P450s. 42
Figure 1.12. Diagrammatic representation of the P450 catalytic cycle. 45
Figure 1.13. Proposed mechanism for hydrogen extraction during substrate catalysis by P450s. 46
Figure 1.14. The P450 heme iron 3d orbital arrangement for the low spin and high spin ferric forms. 47
Figure 1.15. Chemical structures of riboflavin and its derivatives FMN and FAD. 51
Figure 1.16. Ribbon cartoon of the *E. coli* ferredoxin reductase structure. 53
Figure 1.17. Different types of naturally occurring iron-sulphur clusters. 55
Figure 1.18. Chemical structures of A) econazole and B) clotrimazole. 56
Figure 3.1. PCR verification of the orientation of the kanamycin resistance cassette with respect to the *cj1411c* gene in pGEM*/cj1411c::kan* constructs.

Figure 3.2. Schematic diagram of the arrangement of the *cj1411c* gene with the inserted kanamycin resistance cassette.

Figure 3.3. PCR verification of homologous recombination event of the pGEM*/cj1411c::kan* plasmids in PE2 and PE5.

Figure 3.4. Schematic diagram illustrating the cloning strategy used to generate the pCJ0223::cat/cj1411c plasmid.

Figure 3.5. Agarose gel electrophoresis of colony PCR products from *cj1411c* complement strains PE3 and PE6.

Figure 3.6. Reactivity of *C. jejuni* proteins with preabsorbed CYP172A1 antiserum.

Figure 3.7. The role of *cj1411c* insertional knockout and complementation to 24 hour time course growth assays of *C. jejuni* PE1-PE6 strains.

Figure 3.8. Effect of *cj1411c* insertional knockout and subsequent complementation on AAG activity in *C. jejuni*.

Figure 3.9. Transmission electron micrographs investigating the effects of *cj1411c* insertional knockout in *C. jejuni* PE1 and PE2 strains.

Figure 3.10. Transmission electron micrographs investigating the effects of *cj1411c* insertional knockout in *C. jejuni* PE4 and PE5 strains.

Figure 3.11. Effect of *cj1411c* insertional knockout on flagellar-mediated motility of *C. jejuni* 11168H and 81-176 strains.

Figure 3.12. The role of CYP172A1 in mediating polymyxin B sensitivity in *C. jejuni* 11168H and 81-176 strains to polymyxin B containing agar plates.

Figure 3.13. The role of CYP172A1 in mediating azole sensitivity of *C. jejuni* PE1 and PE2 strains.

Figure 3.14. Concentration dependence of econazole mediated bacteriostasis in *C. jejuni* PE1 and PE2.

Figure 3.15. Effect of *cj1411c* insertional knockout and complementation on LOS mobility in *C. jejuni* 11168H.

Figure 3.16. Effect of *cj1411c* insertional knockout and complementation on CPS structures in *C. jejuni* 11168H and 81-176.

Figure 3.17. Alcian Blue staining of *C. jejuni* 11168H P and G strains and their respective ΔkpsM mutants.

Figure 3.18. Averaged FT-IR spectra of *Campylobacter* species and *C. jejuni* strains.

Figure 3.19. PC-DFA ordination plot of *Campylobacter* species and *C. jejuni* strains.

Figure 3.20. Averaged FT-IR spectra of *C. jejuni* strains.

Figure 3.21. PC-DFA ordination plots of three biological replicas of *C. jejuni* strains.
PCA ordination plot showing clusters obtained from the analysis of selected SERS data obtained in *C. jejuni* strains.

Restriction digests of pET15b/cj1411c plasmid construct.

Time dependent production of CYP172A1.

Solubility of CYP172A1 produced in *E. coli* HMS174 (DE3).

Effect of induction growth temperature on the solubility of CYP172A1 produced in *E. coli* HMS174 (DE3).

Production and Ni-NTA purification of CYP172A1 from *E. coli* HMS174 (DE3).

Size exclusion purification of CYP172A1.

Spectroscopic characterisation of ferric, ferrous, ferrous/CO-bound and ferric/NO-bound CYP172A1.

Pyridine hemochromagen complex formation of CYP172A1.

EPR spectrum of native CYP172A1.

Analysis of the stability of heme binding and tertiary structure in CYP172A1 on exposure to increasing concentrations of GdnHCl.


MALLS analysis of native CYP172A1.

Thermofluor assay for CYP172A1 using the JBS buffer screen.

Thermofluor assay for CYP172A1 in Na/K phosphate buffer with added JBS additives.

Thermofluor assay for 1-PIM bound CYP172A1 in Na/K phosphate buffer with JBS additives.

SDS-PAGE analysis of S200 fraction and MALLS analysis of CYP172A1 purified in the presence of 100 mM BME.

Thrombin cleavage of the His6 purification tag from CYP172A1.

Key features used by XstalPred to determine the likelihood of CYP172A1 crystallisation.

Diagrammatic representation of CYP172A1 and truncated mutant proteins.

NdeI digestion of mutagenised pET15b/cj1411c.

SDS-PAGE and Western Blot of truncated CYP172A1 mutants.

CYP172A1 E190H purification using size exclusion gel chromatography.

Examination of crystalline material from trials of CYP172A1 in the presence and absence of 1-PIM.

Representation of a homology model structure of CYP172A1.

Optical titration of CYP172A1 with econazole.

Optical titration of CYP172A1 with miconazole.

Chemical structures of A) econazole and B) miconazole.

Optical titration of CYP172A1 with clotrimazole.

Optical titration of CYP172A1 with sodium cyanide.

Optical titration of CYP172A1 with imidazole.

Chemical structures of 1-, 2- and 4-PIMs.
Figure 5.8. Optical titration of CYP172A1 with 1-PIM. 203
Figure 5.9. Optical titration of CYP172A1 with 4-PIM. 204
Figure 5.10. Difference spectra generated by selected library compound additions to CYP172A1. 210
Figure 5.11. Optical titration of CYP172A1 with 204073. 212
Figure 5.12. Optical titration of CYP172A1 with 204074. 213
Figure 5.13. Chemical structures of compounds 204073 and 204074. 214
Figure 5.14. Optical titration of CYP172A1 with 213071. 216
Figure 5.15. Chemical structures of 204073 and 213071. 215
Figure 5.16. Optical titration of CYP172A1 with 204591. 219
Figure 5.17. EPR spectra of ligand-bound and ligand-free forms of CYP172A1. 222
Figure 5.18. EPR spectra of ligand-bound forms of CYP172A1. 223
Figure 5.19. Chemical structure of the 204591 compound. 223
Figure 5.20. Redox potentiometry of substrate-free CYP172A1. 225
Figure 5.21. Redox potentiometry of 204073-bound CYP172A1. 227
Figure 5.22. Trapping of the Fe(II)CO complex of CYP172A1 using sodium dithionite. 229
Figure 5.23. Formation and collapse of Fe(II)CO complex by CYP172A1 using the E. coli FldR/FldA redox partners. 232
Figure 5.24. Formation and collapse of Fe(II)CO complex of CYP172A1 in presence of 204073 substrate using the E. coli FldR/FldA redox partners. 233
Figure 5.25. Formation the of Fe(II)CO complex of CYP172A1 in presence and absence of 24073 using the E. coli FldR and SpFdx. 236
Figure 5.26. LC-MS analysis of 213071 turnover by CYP172A1. 239
Figure 5.27. LC/MS trace of 213071 turnover reaction by CYP172A1. 241

List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1</td>
<td>Antimicrobial resistance in C. jejuni.</td>
<td>25</td>
</tr>
<tr>
<td>Table 1.2</td>
<td>Summary of solved CPS structures of C. jejuni HS Penner type strains.</td>
<td>33</td>
</tr>
<tr>
<td>Table 1.3</td>
<td>Cytochrome P450 redox systems.</td>
<td>50</td>
</tr>
<tr>
<td>Table 2.1</td>
<td>Antibiotics used in the course of this study.</td>
<td>61</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>List of Campylobacter jejuni NCTC 11168H and 81-176 strains and mutants used in this study.</td>
<td>62</td>
</tr>
<tr>
<td>Table 2.3</td>
<td>PCR program used to amplify target DNA.</td>
<td>64</td>
</tr>
<tr>
<td>Table 2.4</td>
<td>List of antimicrobials tested against C. jejuni.</td>
<td>69</td>
</tr>
<tr>
<td>Table 2.5</td>
<td>Escherichia coli strains used in this study.</td>
<td>74</td>
</tr>
<tr>
<td>Table 2.6</td>
<td>PCR cycling parameters for SDM.</td>
<td>77</td>
</tr>
<tr>
<td>Table 2.7</td>
<td>Cell sample preparation for SDS-PAGE analysis.</td>
<td>78</td>
</tr>
<tr>
<td>Table 2.8</td>
<td>List of mediators used in redox potentiometry.</td>
<td>83</td>
</tr>
<tr>
<td>Table 2.9</td>
<td>List of buffers present in JBS buffer screen and their respective pH values*.</td>
<td>87</td>
</tr>
</tbody>
</table>
Table 2.10. List of additives present in JBS buffer screen and their concentrations.
Table 2.11. Elution profile used to separate 213071 turn over reaction products.
Table 3.1. Viable counts of *C. jejuni* 11168H and 81-176 cells over 16 hours of growth.
Table 3.2. The role of CYP172A1 in conferring antimicrobial sensitivity in *C. jejuni* 11168H and 81-176 strains.
Table 3.3. Summary of azole binding to CYP172A1.
Table 3.4. Summary of type II compounds binding to CYP172A1.
Table 3.5. List of ligands tested for HS formation in CYP172A1.
Table 3.6. ChemDiv substrate and inhibitor compounds and their affinity for CYP172A1.
Table 3.7. Comparison of different redox partners with respect to rate constants for P450 complex formation and collapse for CYP172A1.

List of abbreviations

°C  degree Celsius
δ-ALA  5-aminolevulinic acid
Ω  ohm
µF  micro-Farad
µg  microgram
µl  microlitre
µM  micro molar
1/2/4-PIM  1/2/4-phenylimidazole
AAG  autoagglutination
Amp  ampicillin
BME  β-mercaptoethanol
Bp  base pairs
Carb  carbenicillin
CBA  Columbia Blood Agar
CD  Circular Dichroism
Cfu ml⁻¹  colony forming unit per ml
Chpl  chloramphenicol
CPS  capsular polysaccharide
CO  carbon monoxide
DFA  discriminant function analysis
DMSO  dimethyl sulfoxide
DNA  deoxyribonucleic acid
dNTP  deoxyribonucleic triphosphate
DTT  dithiothreitol
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EB</td>
<td>elution buffer</td>
</tr>
<tr>
<td>EM</td>
<td>electron microscopy</td>
</tr>
<tr>
<td>EPR</td>
<td>Electron Paramagnetic Resonance</td>
</tr>
<tr>
<td>FAD</td>
<td>flavin adenine dinucleotide</td>
</tr>
<tr>
<td>Fdx</td>
<td>ferredoxin</td>
</tr>
<tr>
<td>Fe-S</td>
<td>Iron sulphur</td>
</tr>
<tr>
<td>FldR/FldA</td>
<td><em>E. coli</em> flavodoxin reductase/flavodoxin</td>
</tr>
<tr>
<td>FMN</td>
<td>flavin mononucleotide</td>
</tr>
<tr>
<td>FT-IR</td>
<td>Fourier transform-infrared (spectroscopy)</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>GdnHCl</td>
<td>guanidine hydrochloride</td>
</tr>
<tr>
<td>HS</td>
<td>high spin</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani growth medium</td>
</tr>
<tr>
<td>LC</td>
<td>liquid chromatography</td>
</tr>
<tr>
<td>LOS</td>
<td>lipooligosaccharide</td>
</tr>
<tr>
<td>LS</td>
<td>low spin</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>Kan</td>
<td>kanamycin</td>
</tr>
<tr>
<td>$K_d$</td>
<td>dissociation constant</td>
</tr>
<tr>
<td>kDa</td>
<td>kilo Dalton</td>
</tr>
<tr>
<td>kg</td>
<td>kilogram</td>
</tr>
<tr>
<td>KP_i</td>
<td>potassium phosphate</td>
</tr>
<tr>
<td>kW</td>
<td>kilo volt</td>
</tr>
<tr>
<td>L</td>
<td>litre</td>
</tr>
<tr>
<td>lb/in$^2$</td>
<td>pounds per square inch</td>
</tr>
<tr>
<td>m</td>
<td>minute</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>MALLS</td>
<td>multiangle laser light scattering</td>
</tr>
<tr>
<td>mg</td>
<td>milligram</td>
</tr>
<tr>
<td>MH</td>
<td>Mueller-Hinton medium</td>
</tr>
<tr>
<td>MIC</td>
<td>minimum inhibitory concentration</td>
</tr>
<tr>
<td>ml</td>
<td>millilitre</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NADPH</td>
<td>nicotinamide adenine dinucleotide phosphate (reduced)</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>NO</td>
<td>nitric oxide</td>
</tr>
<tr>
<td>OD$_{600}$</td>
<td>optical density at 600 nm wavelength</td>
</tr>
<tr>
<td>O/N</td>
<td>overnight</td>
</tr>
<tr>
<td>P450</td>
<td>cytochrome P450</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PCA</td>
<td>principal component analysis</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>Rif</td>
<td>rifampicin</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SERS</td>
<td>surface enhanced Raman spectroscopy</td>
</tr>
<tr>
<td>SOB</td>
<td>Super Optimal Broth</td>
</tr>
<tr>
<td>SOC</td>
<td>Super Optimal broth with Catabolite repression</td>
</tr>
<tr>
<td>t</td>
<td>time</td>
</tr>
<tr>
<td>TB</td>
<td>Terrific Broth medium</td>
</tr>
<tr>
<td>TCEP</td>
<td>tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TE</td>
<td>tris EDTA</td>
</tr>
<tr>
<td>Tet</td>
<td>tetracycline</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>v/v</td>
<td>volume per volume</td>
</tr>
<tr>
<td>w/v</td>
<td>weight per volume</td>
</tr>
<tr>
<td>YT</td>
<td>Yeast tryptone medium</td>
</tr>
</tbody>
</table>
Abstract

The University of Manchester

Peter C. Elliott

Characterisation of the unique *Campylobacter jejuni* cytochrome P450, CYP172A1

June 2013

*Campylobacter jejuni* is a leading cause of food poisoning and according to the World Health Organisation accounts for majority of the 4.5 billion cases of global food poisoning each year. Genome sequencing by Parkhill *et al.* (2000) identified a gene, *cj1411c*, which is thought to encode a lone cytochrome P450, CYP172A1. In this thesis the role of CYP172A1 was studied using *in vivo* and *in vitro* techniques. The genomic location of *cj1411c* is adjacent to the capsular biosynthetic genes. The capsular and P450 genes are conserved in some species of *Campylobacter* and *Helicobacter*, as well as in *Comamonas testosteroni*. Importantly, this work has demonstrated that the P450 gene is expressed in two well characterised laboratory *C. jejuni* strains, 11168H and 81-176. Protein production was disrupted using insertional knockout mutagenesis, which allowed for investigations into the role of the enzyme in the host. Alterations to the observed autoagglutination rate and growth characteristics indicate that CYP172A1 has a role in modifying the bacterial surface. The insertional knockout mutant also resulted in cells which were more susceptible to detergent-like compounds (e.g. polymyxin B and sodium deoxycholate). In a previous report, it was suggested that the loss of the P450 function resulted in bacteria which were “shorter and fatter”, compared to wild type cells, but this thesis could find no evidence of such a phenomenon. CYP172A1 was successfully purified using recombinant expression in *E. coli* to enable biochemical and biophysical characterisation *in vitro*. CYP172A1 contains a typical P450 cysteine thiolate coordination to the heme iron, and exists in a low spin ferric heme state under neutral buffer conditions. The P450 was found to self aggregate, and despite rigorous investigations the cause of this aggregation was not fully established. Despite this issue, CYP172A1 was shown to bind to a wide range of P450 inhibitor-type compounds, with econazole displaying the tightest binding affinity (*K_d* = 100 nM). Identification of substrate-like compounds was achieved using high throughput compound screening, and a number of organic compounds were identified and shown to bind CYP172A1, inducing heme iron absorbance changes typical of either P450 inhibitors or substrates. Optical titrations for these molecules indicated that their CYP172A1 *K_d* values were in the low micromolar range. The catalytic capability of CYP172A1 was successfully demonstrated by providing the P450 with non native redox partners to oxidise one of such substrate-like compound (213071), resulting in the sulfoxidation of this compound.
Declaration

The author declares that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), 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

This thesis is dedicated to the memory of my mum, a person to whom I owe such much and miss every day.

I would like to thank my supervisor, Professor Andrew Munro, for all of his help and guidance throughout my time as his PhD student. Without his support I doubt I would have been able to reach this point. I consider myself very lucky to have been part of such a great group. Although there have been comings and goings at various stages you all made me feel at ease and helped me through the years. I feel very honoured to be able to call myself a member of the “Munlop” team.

Special mentions must go to Dr Kirsty Mclean for her tireless attitude towards improving my understanding and knowledge of both P450s and general laboratory skills. This also extends to the other senior members of the lab. I cannot go without also mentioning Marina Golovanova who as well as be patient in her teachings also permitted me to get away with the most outrageous cheekiness, thus maintaining my sanity (for the most part anyway).

I have enjoyed getting to know the rest of the Molecular Enzymology group within the MIB, all of whom have helped me in one way or another and I hope that many of us will remain friends for many years.

I would also like to thank my co-supervisor Dr Dennis Linton for allowing me the use of his microbiology facilities. I extend my gratitude to the senior members of the Linton group for their patience in attempting to teach an enzymology dog some new tricks. I also wish to thank Deborah Cantu-Watson for keeping my spirits up during the tough days.

Finally I would like to thank my family, I know none of you really know what I have been doing with myself for the last four years, but that has not hampered your desire to see me succeed.
Chapter 1. Introduction

1.1.1. The *Campylobacter* genus and related organisms

Although initially considered a species of the *Vibrio* genus, the organism now known as *Campylobacter fetus* was first isolated from aborted sheep fetuses in 1913 [1]. It was not until 1963 that the *Campylobacter* genus was proposed to incorporate isolates clearly different, especially in the G+C content of their genomic DNA, compared to other *Vibrio* species [2]. The *Campylobacter* genus contains 18 species that are Gram negative, and predominantly microaerobic bacteria that colonise the gastrointestinal tracts of humans and animals. The most commonly reported human pathogens are *C. jejuni* and *C. coli*, but *C. fetus, C. lari* and *C. upsaliensis* are also known to be opportunistic pathogens with the rest of the species thought to be commensals that cause no obvious harm to their host [3].

As well as the *Campylobacter* genus, the Campylobacterales order also contains the *Arcobacter, Helicobacter* and *Sulfurospirillum* genera (Figure 1.1). These taxa belong to the epsilon subdivision of Proteobacteria. In general these bacteria are thermophilic and grow in either aerobic or microaerobic conditions [4]. Typically members of the Campylobacterales order are either pathogenic or commensal in humans and domestic animals. Generally they are all Gram negative, have S-shaped morphologies and do not form spores [5]. The bacteria are unable to utilise carbohydrates as a source of energy, but instead use amino acid catabolism. Perhaps the best known of these species is *Helicobacter pylori*, which colonises the human stomach. Persistent *H. pylori* infection is associated with inflammation causing stomach ulcers and gastric cancers [6]. Some of the species of the Campylobacterales order are known or predicted to express an outer capsule (orange boxes, Figure 1.1).
Chapter 1. Introduction

Figure 1.1. Rectangular cladogram of the selected members of the Campylobacterales order. Alignments are based upon the 16S RNA gene sequences. 16S RNA sequences obtained from GenBank were aligned in ClustalW2, and the tree constructed using FigTree software. Species* that are capable of producing an outer capsule are highlighted (orange). C., Campylobacter; H., Helicobacter; A., Arcobacter; S., Sulfurospirillum.

1.1.2 *Campylobacter jejuni* and its general features

*C. jejuni* was first identified in 1973 and is now known to be the leading bacterial cause of food poisoning throughout the world, affecting more people than *Escherichia. coli* and *Salmonella* combined [5, 6]. The World Health Organization (WHO) estimates that globally there are 4.5 billion cases of diarrhoeal disease each year and, as the lead cause of food poisoning, it is assumed that *C. jejuni* makes up a significant number of those cases [7]. In the USA alone it is estimated that *C. jejuni* infects 2.5 million people annually, ~1% of the population [8]. In 2012, the European Centre for Disease Prevention and Control reported that infection by *C. jejuni* was the most commonly cause of food-borne outbreaks in Europe since 2005, citing over 200,000 cases of zoonosis in 2010 [9].
*C. jejuni* has a spiral shaped and has a single flagellum at each pole (Figure 1.2) [10]. Complete flagella are important in motility in chickens and infection in humans [11, 12]. The bacterium is relatively small in size, 1.0 by 0.5 µm [13], has an optimum growth temperature at 42 °C, is microaerophilic and is unable to grow aerobically or anaerobically. The bacterium usually resides in the intestinal tract of its hosts. Infection in humans is generally self limiting, but further complications can occur (see Section 1.2).

![Image](242x464 to 413x554)  
*Figure 1.2* Scanning electron micrograph of *C. jejuni*. Cells have a spiral morphology with a single flagellum at each pole. Image acquired from Altekruse et al. (1999).

The *C. jejuni* genome is 1.6 Mbp in size [14]. This is relatively small compared to other bacterial pathogens, e.g. enterotoxigenic *E. coli* Strain H10407 (5.1 Mbp) and *Salmonella typhimurium* (4.8 Mbp). However, ~95% of the genome encodes for protein, making it one of the most efficient bacterial genomes [14-16]. Analysis of the genome also revealed that the genes are generally not organised into operons of related genes with the notable exceptions of the lipooligosaccharide (LOS), capsule polysaccharide (CPS) and flagellum biosynthetic clusters. Within these gene clusters there are a number of genes containing intragenic poly G tracts that can mispair during replication. The resultant frame shifting leads to phase variable gene expression, as demonstrated in *H. influenzae* [17]. *C. jejuni* is also naturally competent and able to take up DNA from the environment, which then recombines onto the chromosome. DNA exchange occurs more readily between related species [18].
1.1.3. *C. jejuni* infection sources

Humans come into contact with *C. jejuni* from the handling, preparation and consumption of uncooked chicken meat and, more likely in developing countries, from contaminated water sources. *C. jejuni* is a zoonotic organism that can survive in poultry, pigs, livestock, wild animals and humans, and water can also act as a reservoir for *C. jejuni* (Figure 1.3) [19, 20]. Poultry is the most common source of *C. jejuni* in the developed world, with chickens being the primary host. It is interesting to note that, in avian species, infection by *C. jejuni* is actually low and the bacterium appears to be more commensal rather than being a pathogen. The bacterium does not invade into the epithelia, but instead colonises the gut just above the epithelial cells in the mucus layer. However, in man *C. jejuni* infection is invasive [21, 22]. Why this occurs is not yet understood. One possibility is that this results from differences in temperature the bacteria experience. Avian species have an internal body temperature of 42 °C whereas the human gut temperature is only 37 °C. This change in temperature might affect the production and stability of several key invasive proteins [23].

![Figure 1.3. Transmission routes of *C. jejuni* between hosts. *C. jejuni* is most commonly found in the intestinal tract of poultry. *C. jejuni* can infect humans from contaminated water sources, ill prepared chicken meat or contaminated unpasteurised milk. Adapted from Young et al. (2007).](image-url)
1.2. *C. jejuni* infection and symptoms

It can take as little as $10^2$ *C. jejuni* cells to cause human infection [24], which is much lower than for many other pathogenic bacteria, including *Vibrio cholerae* which requires the ingestion of $10^6$ - $10^9$ cells for infection [25]. *C. jejuni* travels through the stomach into the small intestine, where it rapidly replicates. The infection is generally self-limiting and typically lasts for 7-10 days. Symptoms present approximately three days post consumption, and include bloody diarrhoea, stomach cramps and nausea [3]. Death from infection is rare, but is possible if severe dehydration occurs or if other underlying medical conditions are present [26].

In ~1/1000 cases of *C. jejuni* infection Guillain-Barré syndrome (GBS) or the related Miller Fisher syndrome occurs, which can result in temporary paralysis. The link between GBS following *C. jejuni* infection was first noted in 1982 [27]. GBS occurs due to molecular mimicry between bacterial LOS and human gangliosides. Evidence has indicated that GBS can result following infection of any serotype of *C. jejuni*, but that serotypes HS:19 and HS:41 are the most commonly identified following infection [28]. The likelihood for developing GBS post infection with serostrain HS19 is 1/200, which is 5x more likely compared to the average [29]. In an attempt to clear the infection, the immune system produces antibodies that target the sialic acid on the bacteria, which cross react with host gangliosides damaging the myelin and nerve axons surface, resulting in paralysis. Paralysis is typically temporary and any loss of mobility is fully restored. However, in a small fraction of cases the effects persist for several years [30].

1.2.1. *C. jejuni* adhesion to epithelial cells

Once the bacterium has penetrated the mucus it can interact with the host epithelial cells. *C. jejuni* adheres to these cells to prevent the mechanical action of the mucus removing the bacterium out of the GI tract before replication can occur. There are several proteins thought to play a role in adhesion of *C. jejuni* to epithelia, including CadF, PEB1 and FlpA, all of which target different structures on the host cell [31]. One of the main targets the bacterium recognises is the extracellular matrix component fibronectin. *C. jejuni* produces two fibronectin binding proteins: Campylobacter
adhesion to fibronectin (CadF) and fibronectin-like protein A (FlpA), which mediate binding to epithelial cell lines [32]. These proteins also activate Rac1 and Cdc42 (host cell Rho GTPases), which are essential for entry into the epithelial cell [7]. The PEB1 protein is also important in cellular interactions and mutations that affect PEB1 functionality reduce bacterial adherence to epithelial cells by 50 to 100 fold [33].

1.2.3. C. jejuni invasion of epithelial cells

Typically, invasion is limited to the human host with the bacteria affecting host cell signalling cascades to aid internalisation. Recently, Krause-Gruszczynska et al. (2011) investigated the signalling cascade and the proteins involved in order to propose a method for C. jejuni internalisation (Figure 1.4). They found, through gene knockout studies, that CadF was important in binding to fibronectin in the extracellular matrix. The fibronectin is attached to integrins and attachment to C. jejuni activates the tyrosine kinases focal adhesion kinase (FAK) and Src. Src interacts with PDGRF/EGFR (platelet derived growth factor receptor/epidermal growth factor receptor) located in the cell membrane, which activates the phosphatidylinositol 3-kinase (PI3-K) protein. PI3-K in turn activates the guanine nucleotide exchange factor Vav-2, which can then activate the Rho GTPase Cdc42. Cdc42 activation then leads to cellular rearrangement and C. jejuni engulfment. Mutations to each of the proposed components of this signalling pathway resulted in a reduction of C. jejuni internalisation [7]. Another set of important proteins, that are only secreted upon contact with the epithelium, are the Cia (Campylobacter invasion antigen) proteins [34]. The function of the Cia proteins is not presently known, but gene knockout studies have indicated that mutants lacking the CiaB protein fail to cross the epithelial cell membrane [35]. The Cia proteins are delivered through the flagella, using a type III secretion pathway, and so it is assumed that the flagella must penetrate the epithelial membrane [36]. Once internalised into host cells (via a phagosome), the bacterium is able to avoid degradation by preventing fusion with lysosomes using a currently undefined method [7].
Figure 1.4. Model of *C. jejuni* internalisation via the activation of the Rho GTPase Cdc42. CadF binds to integrin connected fibronectin. Clustering of integrins leads to the activation and phosphorylation of the tyrosine kinase FAK. Phosphorylation of FAK and Src starts an internal signal cascade as indicated. Integrin activation leads to the phosphorylation and transactivation of the PDGFR/EDGFR. PI3-K and Vav-2 activation then occurs inducing the activation of Cdc42. Cdc42 activation leads to the rearrangement of microtubules and F-actin, allowing internalisation of *C. jejuni*. Along with CadF the flagellum provides an alternative entry method for *C. jejuni*. Image acquired from Krause-Gruszczynska et al. (2011).

1.3. Antibiotic resistance in *Campylobacter* species

As *Campylobacter* is a human and veterinary pathogen, a range of antimicrobials have been developed to treat infection. Although the disease is generally self limiting, there are cases where drug treatment is advised, including those with immune deficiencies (e.g. HIV patients), pregnant women or those with long lasting infection [26]. There are four techniques recognised by the National Committee for Clinical Laboratory Standards that can be used to determine the MIC (minimal inhibitory concentration) values of antimicrobials against *Campylobacter*. These are agar dilution, microbroth dilution, E testing and disc diffusion. To date, no single technique has been chosen as the international standard [37]. Some studies have shown that the results from different methods have been comparable for some antimicrobials [37].
The resistance to antimicrobials by *Campylobacter*, especially *C. jejuni* has increased over the last three decades due to misuse. Fluoroquinolones are very effective at treating infection, as they prevent DNA replication and were used in livestock and poultry by the FDA (Food and Drug Administration, USA) in 1996 in an attempt to stem the infection at source [26, 38]. This approach proved to be ineffective, although whether this was due to inappropriate dosage is not known. As a result of treatments with the drug, the bacterium evolved resistance mechanisms. In 1990 fluoroquinolone resistant *Campylobacter* were not detected in the US, but by 1999 18 % of isolates were resistant to this antibiotic [1]. This trend was mirrored across Europe and the continued use of these drugs only provided further selective pressures that increased the amount of resistance in the organism. Resistance to fluoroquinolones was first identified as a single point mutation in the gene encoding for the A subunit of DNA gyrase [39]. The location of this mutation is in the adeptly named quinolone resistance determining region. The most frequently reported mutation is the Thr86Ile, which increases the cell’s ciprofloxacin MIC values from 16 to 64 µg/ml [40, 41]. Horizontal gene transfer allowed for the spread of resistance among strains [42]. Antimicrobial resistant *Campylobacter* can prolong the illness duration by 2-3 days on average [37]. In Australia these fluoroquinolones have not been extensively used and so *Campylobacter* remains largely susceptible to their action.

Worldwide resistance to other antimicrobials, including tetracycline, macrolides and chloramphenicol, has been observed. The modes of resistance are different for the majority of these agents (Table 1.1). More worrying is the emergence of multidrug resistant (MDR) strains due to the activity of efflux pumps that prevent the internal concentration of antimicrobials from accumulating to lethal levels. Recently, it was also shown the CmeABC efflux system plays a role in fluoroquinolone resistance, which highlights the bacterium’s ability to create several resistance mechanisms to improve its survival chances when challenged with antimicrobial agents [41]. As there are few effective antimicrobials remaining, it is important to produce new compounds that mediate killing of *Campylobacter* and to manage their use to prolong their efficacy for as long as possible.
Chapter 1. Introduction

### Antimicrobial Resistance mechanisms

<table>
<thead>
<tr>
<th>Antimicrobial</th>
<th>Resistance mechanisms</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aminoglycoside</td>
<td>Aminoglycoside modification by a 3’-aminoglycoside phosphotransferase.</td>
<td>[43]</td>
</tr>
<tr>
<td>β-lactams</td>
<td>Enzymatic alteration, reduced uptake, efflux</td>
<td>[44]</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>Modification of compound by chloramphenicol acetyltransferase.</td>
<td>[45]</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>Ribosomal protection by TetO and/or efflux pump.</td>
<td>[44]</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>3 point mutations on 23S rRNA.</td>
<td>[44]</td>
</tr>
<tr>
<td>Fluoroquinolones</td>
<td>Point mutation in the Quinolone Resistance Determining Region (QRDR) of the gyrA subunit-encoding gene and efflux pumps.</td>
<td>[46, 47]</td>
</tr>
</tbody>
</table>

**Table 1.1. Antimicrobial resistance in C. jejuni.** Several antimicrobials are no longer effective against various C. jejuni stains as a consequence of diverse bacterial resistance mechanisms.

### 1.4. Surface structures of C. jejuni.

The surface of the human pathogen C. jejuni plays an important role in maintaining cell integrity, survival in harsh environments, avoidance of host immune systems, colonisation and virulence [48]. The surface is predominantly made up of three components, the flagellum, the lipooligosaccharide (LOS) and the capsule polysaccharide (CPS) (Figure 1.5), and it is these three features that will be the focus of subsequent sections.

#### 1.4.1. The Flagellum

Flagella mediated motility is considered vital to survival in the diverse environments C. jejuni is exposed to (e.g. water bodies, intestine *etc*). Motility is required for movement through intestinal mucus and for invasion of epithelium, and it is controlled by aerotaxis and chemotaxis systems. C. jejuni is actively motile towards amino acids found in the chicken intestine [49]. It is also known that the flagellum is involved in secreting proteins that might affect virulence (e.g. Cia and FlaC proteins) [50]. As is the case with the flagellum from other bacteria, it is made up of three parts: a membrane embedded motor, the hook and the outer filament [51]. The flagellar filament is made up of two
proteins, FlaA and FlaB. FlaA is the major constituent, but both proteins are required for complete flagellum assembly and full motility [52].

![Diagram of C. jejuni outer surface structures](image)

**Figure 1.5. Diagrammatic representation of the outer surface structures of C. jejuni.** The structures of the flagellum, LOS and CPS are located in/on the outer membrane of the bacterium. The LOS contains an inner (purple square) and outer (green square) sugar composition attached to a lipid A moiety. The CPS is composed of repeating di/trisaccharide linkages. The flagellum filament/tail predominantly contains flagellin protein and O-linked glycans (yellow circle). Adapted from Young *et al.* (2007).

The flagellin has two interesting structural features. The first is a reduced capacity to bind to toll-like receptors (in particular TLR5) and the second is the presence of high levels of glycosylation of the flagellin protein [50]. Toll-like binding receptors bind conserved regions of bacterial flagellin and are used by the innate immune system as a way to immobilise foreign cells [53]. The absence of these recognition sites will assist *C. jejuni* in avoiding the innate immune system. The flagellin protein is glycosylated with pseudaminic acid and its derivatives via O-linkages. Nine genes have been identified in *C. jejuni* 81-176 (which has the simplest glycosylation locus) as associated with the biosynthesis and transfer of pseudaminic acid [54]. Analysis of other *Campylobacter* genomes has revealed that these glycosylation genes are among the most variable, and give rise to a multitude of different glycans [50]. Flagellar interactions are important for *Campylobacter* in aiding the formation of microcolonies, whereby groups of cells clump together around epithelial cells prior to invasion. Loss of flagellar glycan
structures results in a 60 % reduction in invasion and an 80 % reduction in adherence [50].

As well as C. jejuni’s ability to perform O-linked glycosylation, the bacterium is capable of the N-linked glycosylation of proteins. O-linked glycosylation is limited to the flagellin protein, whereas there are many N-linked glycosylation target proteins. The C. jejuni N-linked glycosylation pathway was the first example of this type of glycosylation in bacteria [48]. The pathway involves biosynthesis of a heptasaccharide glycan, which is transferred onto proteins with an asparagine residue within a consensus sequence of Asn-X-Ser/Thr where X can be any residue except for a proline [55, 56].

N-linked protein glycosylation occurs in the periplasm and is mediated by an inner membrane located oligosaccharide transferase PglB [56]. A total of 12 genes have been identified in this system termed as the pgl (protein glycosylation) locus, pglA-K, which all mediate the biosynthesis and attachment of the glycan. The entire locus has been cloned into E. coli and glycosylation of two target proteins was observed, clearly proving that these genes are necessary and sufficient for N-linked glycosylation [57]. PglB is perhaps the best characterised protein, and its function is to transfer the glycan from the undecaprenyl pyrophosphate carrier to the target protein [56]. Studies which either knocked out the pglB gene or introduced a mutation in a highly conserved region resulted in a loss of N-linked glycosylation [57, 58].

The latest model of the C. jejuni pathway highlights how the glycosylation is achieved (Figure 1.6). Initially PglF, PglE and PglD convert the GalNAc linked uridine diphosphate (UDP) to bacillosamine (Bac) linked UDP. Then PglC transfers the Bac onto undecanprenyl-pyrophosphate (Und-PP), which is anchored in the membrane. PglA attaches a GalNAc to the Bac and PglJ attaches a second GalNAc onto the first GalNac. PglH attaches three GalNAc residues in a sequential order before PglI attaches a branched glucose to create the full length glycan. PglK (or WlaB) then transports glycan across the membrane into the periplasm. Finally, PglB removes the glycan from the Und-PP anchor and attaches it to a target protein.
Figure 1.6. Representation of the current model of the Pgl pathway in *C. jejuni*. The production of the heptasaccharide glycan begins with the generation of Bac from GlcNAc. Five GlcNAc’s are sequentially added as is a branched glucose. The glycan is flipped across the membrane into the periplasm and translocated onto the target protein. GlcNAc, N-acetylated glucose; Bac, bacillosamine; GalNAc, N-acetylated galactose; UDP, uridine diphosphate; Und-PP, undecaprenyl-pyrophosphate. Image acquired from Weerapana and Imperiali (2006).

1.4.2. *C. jejuni* lipooligosaccharide (LOS)
The LOS forms an integral part of the cell outer membrane and consists of a lipid A moiety modified with an oligosaccharide structure. Modulation of this oligosaccharide layer can affect interactions with the immune systems and, cell to cell interactions, and also molecular mimicry of the host gangliosides, which can lead to GBS [29, 59, 60].

The oligosaccharide portion of the LOS is comprised of two parts; the more conserved inner core and the more variable outer core. The structure of the *C. jejuni* 11168 LOS core was first reported by Oldfield *et al.*, and later revised (Figure 1.7) [60, 61]. The structure of the NCTC 11168 outer core structurally resembles human gangliosides GM1a and GM2, and it is the presence of the sialic acid which is the key factor that can lead to GBS [62].
Figure 1.7. Sugar composition of the complete LOS core of *C. jejuni* 11168. GM1a and GM2a are carbohydrate structures analogous to human gangliosides. The inner core (black) and outer core (red) of the LOS is displayed along with the positions of additional attachments. The oligosaccharide attaches to lipid A via the Kdo substituent. Kdo, 3-deoxy-D-manno-oct-2-ulopyranosonic acid; Hep, L-glycero-D-mannoheptose; PEtn, pyrophosphorylethanolamine; Glc, glucose; Gal, galactose; Neu5Ac, N-acetylneuraminic acid (sialic acid); GalNAc, N-acetyl-D-galactosamine. Image adapted from St Michael *et al.* (2002).

The LOS structure is highly changeable due to variable expression of the genes involved in its biosynthesis, which suggests a possible mechanism/strategy for avoiding host immune systems. The variation of LOS structures observed in different *C. jejuni* species arises from alterations in the genes present in the LOS loci as well as sequence variation, which alter the levels of gene expression [63]. Phase variation due to slipped strand mispairing at intragenetic homopolymeric tracts also plays an important role in which genes are expressed, and nine genes within the LOS locus have been identified as containing these tracts [14]. One such phase variable gene, *wlaN*, encodes a galactosyltransferase and is responsible for transferring the terminal β 1→3 Gal residue (Figure 1.7). When this gene contains an intragenic homopolymeric tract of nine G’s a stop codon is introduced, converting the GM1a structure to the GM2. A tract length of eight G’s reintroduces the GM1a structure, by removal of the stop codon [59]. The importance of an intact LOS has been demonstrated by genetic manipulations. The mutation of the *waaC* gene results in a truncated LOS, lacking any sugar beyond the Kdo substituent. This mutation results in the cell being less invasive, which has been suggested to be a result of decreased membrane stability [64]. This again highlights that the LOS is an important factor in cell survival and virulence.
1.4.3. *C. jejuni* capsule polysaccharide (CPS) locus

Until recent analysis of the genome identifying the cps gene cluster, it was thought that, along with the LOS, *C. jejuni* produced a lipopolysaccharide-type glycolipid, a high molecular weight LOS with O-linked attachment [65]. Analysis of the genome of *C. jejuni* NCTC 11168 revealed that the strain contained a type II/III capsule locus [66], similar in overall arrangement to those been identified in *E. coli* K1 and *Neisseria meningitidis* [67]. The identification of *kps* genes within the genome means that is now commonly accepted that *C. jejuni* produces CPS [68]. The genetic region that encodes for the CPS is composed of genes involved in capsule biosynthesis and genes involved in capsule translocation to the cell surface (*kpsC-F, kpsM, kpsS* and *kpsT*). The *kps* genes are found in only 4 of the 18 *Campylobacter* species (*C. jejuni, C. coli, C. lari and C. upsaliensis*), and show high levels of homology. However, the biosynthetic genes display a large amount of variability, with similar genes either absent or structurally altered between *Campylobacter* species [69]. The only exceptions to this concern three genes involved in the synthesis of the heptose sugars, *hddC, gmhA* and *hddC* [70].

Unlike the LOS, the attachment of the CPS to the cell is not via lipid A, but instead is proposed to involve a 1,2-dipalmitoyl-glycerophosphate anchor whereby the polysaccharide attaches to the phosphate group [71]. The *C. jejuni* capsule has two key unusual features compared to other encapsulating bacteria, a) the conformations of the heptose sugars (*alto, ido, gulo and talo*) and b) O-methyl phosphoramidate (MeOPN) moieties [70]. The CPS is on the outer most surface of the bacterium and has been linked to serum resistance, adherence and invasion of epithelia, and virulence [19].

1.4.4. *C. jejuni* CPS structural variability and serotyping

Recently it has been discovered that serodeterminant of the Penner serotyping scheme of *C. jejuni* is the CPS [66]. Multiple methods of typing or identifying different strains of *Campylobacter* have been used, but generally fall into two categories. These are ones
involving the use of cell structures and ones involving genomic sequencing [72]. The
Penner serotype system is a method that is perhaps the most widely accepted and
employed to identify Campylobacter strains and uses the passive haemagglutination
towards Campylobacter to identify the serotype [73]. At present, 47 different serotypes
have been identified in C. jejuni, which highlights the amount of CPS structural
variation [70, 72].

To date several CPS structures have been solved, including those for C. jejuni strains
11168 (Heat Stable or HS:2) and 81-176 (HS:23/HS:36). This was achieved using a
range of techniques, including gas chromatography, mass spectrometry and NMR
(Table 1.2.) [61, 68, 74, 75]. The CPS backbone is made up of heptose sugars of two or
three repeats. To this core structure, several labile groups can be added, including
MeOPN, glycerol and ethanol substituents. This vastly increases CPS complexity and is
most likely a method that the bacterium uses to avoid host immune systems. The
variation in the solved structures is attributed to multiple mechanisms of their alteration,
including phase variation [14, 66, 76]

It was recently shown that the cationic dye Alcian Blue is able to bind to the CPS of
several strains of C. jejuni and can be used to detect the capsule of the cell surface in
electron micrographs, highlighting the capsule structure around the cell [66, 77]. The
dye is thought to work by binding to negatively charged constituents of the CPS,
including polar carbonyl (C=O) groups.

Because the CPS is so exposed to the bacterium’s surroundings, it has the potential to be
used to produce a subunit vaccine. The benefit of using just a cellular subunit, rather
than an attenuated strain, is the reduction in potential complications upon vaccination.
Other encapsulated bacteria have had successful vaccines created in this way, e.g. H.
influenzae [70]. Three conjugate vaccines have been created for HS4, HS15 and
HS23/36 [78, 79]. The CPS is purified and attached onto a mutated diphtheria toxin
subunit, via a deamination reaction, to ensure it is recognised by the host, as
carbohydrates on their own do not produce a T cell mediated immune response, which is
vital for stimulating immunological memory [80]. In mouse studies the three vaccines
all elevated IgG levels, which indicated that the conjugate was being recognised. The
HS23/36 vaccine was also given to non primate monkeys and the vaccination prevented diarrhoea in all of the monkeys, compared to the control group, 70% of which suffered from diarrhoea [78]. The major challenge faced in developing a CPS-based vaccine is the large number of different serotypes, reflecting CPS structural variation. Because of this a multivalent vaccine will be required, although it will most likely not need to include every potential serotype, as different serotypes appear to have distinct geographical distributions [78].
### Table 1.2. Summary of solved CPS structures of C. jejuni HS Penner type strains

The core backbone repeat is given along with any further side chain additions and the positions at which these additions occur on the core. P, phosphate; Gal, galactose; Gro, glycerol; Me, methyl; Hep, heptose; Rib, ribose; GalNAc, N-acetylgalactosamine; GlcA, glucuronic acid; NGro, aminoglycerol; GlcNAc, N-acetylglucosamine; Ara, arabinose; Alt, altrose; Fuc, fucose.

<table>
<thead>
<tr>
<th>Serotype</th>
<th>CPS core structure</th>
<th>Side Chains</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HS:1</td>
<td>[→4)-α-D-Galp-(1→2)-(R)-Gro(1-P→]</td>
<td>-6-O-Me-β-1-glc-Hepp (C3 GlcA)</td>
<td>[68]</td>
</tr>
<tr>
<td>HS:2</td>
<td>[→2)-β-D-Ribf-(1→5)-β-D-GlcpNAc-(1→4)-α-D-GlcpA6(NGro)-(1→]</td>
<td>-MeOPN (O-2 or O-7 6d-ido-Hep)</td>
<td>[61]</td>
</tr>
<tr>
<td>HS:4</td>
<td>[→3)-6d-D-ido-Hepp-(1→4)-β-D-GlcpNAc-(1→]</td>
<td>-α-1-sorbofuranose (C2 GlcA), -MeOPN (C4 GlcNAc)</td>
<td>[74]</td>
</tr>
<tr>
<td>HS:15</td>
<td>[→3)-α-L-Araf-(1→3)-6d-α-1-gulo-Hepp-(1→]</td>
<td>-</td>
<td>[79]</td>
</tr>
<tr>
<td>HS:19</td>
<td>[→4)-β-D-GlcA6NGro-(1→3)-β-D-GlcpNAc-(1→]</td>
<td>-</td>
<td>[75]</td>
</tr>
<tr>
<td>HS:23/HS:36</td>
<td>[→3)-β-D-GlcpNAc-(1→3)-α-D-Gal-(1→2)-D-glycero-α-D-altro-D-Hep-(1→]</td>
<td>-3-O-Me-D-glycero-α-D-altro-D-Hep-(1→)</td>
<td>[81]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-6d-α-D-altro-Hep (1→)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-6d-3-O-Me-α-D-altro-Hep (1→)</td>
<td></td>
</tr>
<tr>
<td>HS:41</td>
<td>[→2)-β-L-Araf-(1→2)-β-D-6d-altro-Heppf-(1→2)-β-L-6d-Altf-(1→]</td>
<td>-α-Xlu (C2 β-6d-D-manno-heptopyranose), -β-Xlu (C4 β-6d-D-manno-heptopyranose)</td>
<td>[82]</td>
</tr>
<tr>
<td>HS:53</td>
<td>[→3)-β-6d-D-manno-heptopyranose-(1→3)-α-6d-D-manno-heptopyranose-(1→3)-α-6d-D-manno-heptopyranose-(1-P→]</td>
<td>-</td>
<td>[83]</td>
</tr>
</tbody>
</table>
1.5. The presence of CYP172A1 in Campylobacter jejuni.

Parkhill et al., (2000) sequenced the C. jejuni genome and their findings revealed the presence of a single cytochrome P450 (P450 or CYP) enzyme named CYP172A1, which is encoded by the \( cj1411c \) gene. The gene is located adjacent to the \( kps \) gene cluster, involved in the biosynthesis and transport of the bacterium’s CPS (Figure 1.8.). \( Cj1411c \) is surrounded by two genes of unknown function (\( cj1410c \) and \( cj1412c \)), but both encode possible membrane associated proteins. Upstream of the P450 gene, but on the other DNA strand, are two genes, \( filL \) and \( acpS \), which are putative genes thought to be involved in flagellar functionality and LOS biosynthesis, respectively [84].

![Figure 1.8. Location of \( cj1411c \) in the C. jejuni genome.](image)

The P450 gene is found in other bacteria including other members of the Campylobacter genus (C. coli, C. lari and C. upsaliensis), three Helicobacter strains and also Comamonas testosterone (Figure 1.9). Comparisons of the gene products surrounding the P450 indicate equally strong amino acid similarity, with the exception of C. testosterone, when compared to those in C. jejuni. The genetic location is also highly conserved, with \( cj1411c \) being located near to the \( kps \), \( filL \) and \( acpS \) genes, in all the bacteria that contain the P450 gene. The P450 gene is not found in all species of Campylobacter and Helicobacter (e.g. C. hominis, C. rectus and H. pylori), which might suggest that the gene was lost in those species that did not require the function of the protein in order to occupy their various niches, but that the protein plays an important role in those organisms that still retain it.
Figure 1.9. Diagrammatic representation of the *C. jejuni* P450 gene (*cj1411c*) and the surrounding genes aligned with orthologous genes from related bacteria. Numbers below gene arrows correspond to % amino acid similarity of the corresponding orthologous gene product from *C. jejuni*. The size of the arrow is proportional to gene size. P450 gene (red), Flagellar assembly gene (green), lipid assembly gene (orange), capsule transport gene (lilac), *galE* (yellow) and gene of unknown function (grey).
1.6. History of the cytochrome P450

In 1955, the first reports of a cytochrome P450 were made independently by Axelrod [85] and by Brodie et al., who identified that an enzymatic system from liver microsomes could oxidize xenobiotics [86]. Later, in 1958, Klingenberg and Garfinkel separately identified a red pigment from liver microsomes which produced a strong spectroscopic absorbance at 450 nm, after the addition of the reducing agent sodium dithionite and on exposure to CO [87, 88]. It is from this discovery that cytochrome P450 (Pigment 450 nm, or P450) obtained its name. To date, the P450s are the only enzyme class that produce such a characteristic peak [89].

P450’s contain a prosthetic group (a protoporphyrin ring with an iron atom at its centre), which is non covalently bound to the protein (apart from in some members of the CYP4 fatty acid hydroxylase family, who contain an ester bond between a heme methyl and either a Glu or Asp carboxylate [90]). The heme is situated near the centre of the P450. The iron atom is ferric in the protein resting state, and is coordinated by up to six ligands; four from pyrole nitrogens in the porphyrin ring in the heme plane, one from a conserved cysteine residue, and the final coordination position (located above the iron atom) is either a water molecule (in the resting state) or dioxygen in the activated (ferrous) form of the protein. The water is weakly bound and can be displaced (e.g. by substrate) to leave a pentacoordinated form.

The binding of the heme to the P450 means that in a UV-Visible spectrum there are commonly five distinct features, and these are known as the alpha (α), beta (β), gamma (γ) and delta (δ) peaks, and finally the protein (aromatic amino acid) absorbance at 280 nm. With the exception of the 280 nm peak, these spectral features arise from the heme. The α and β peaks are typically observed between 500-600 nm, the γ peak (also known as the Soret peak, as the largest of these features) is seen at ~418 ± 3 nm for the low spin ferric heme of a P450, and the δ peak is at ~360 nm. The values of the wavelengths observed are not fixed and can be affected by a number of different factors, including pH and hydrostatic pressure [91, 92]. The absorption bands are also affected by changing environment of the heme, commonly by the introduction of axial ligands different to those found in the protein’s resting state. Under native conditions the ferric...
P450 heme iron is usually in a low spin (LS) state – reflecting the organization of the electrons in the heme iron 3d orbitals - with absorbance peaks in the wavelength regions mentioned above. The sixth axial ligand to the heme is typically a water molecule in the resting state. If a substrate or stronger heme iron ligand is added and can gain access to the heme, then an interaction may occur that alters the coordination and/or spin state of the heme iron. The interaction can either be classed as a type I (substrate-like) or type II (inhibitor-like). A type I interaction causes a blue shift to the Soret peak, inducing a move of the peak to a shorter wavelength, typically ~390 nm. In doing so the ferric heme iron undergoes a shift from a LS to a HS (high spin) state, which occurs as the substrate displaces the coordinating water from the sixth ligand position as it binds in the P450 active site [93]. Ligands that produce a type I shift of the Soret peak are almost exclusively substrate or substrate-like molecules.

In type II binding, a red shift occurs to the Soret peak, moving it to a longer wavelength. A type II shift is caused by the binding of an inhibitor that ligates directly to the heme iron. Typically, an inhibitor contains an atom which has unpaired electrons, which enables interactions with the heme iron. This atom competes with the water present to directly bind to the heme iron in the 6th axial position. The replacement of the water with a stronger binding ligand stabilizes the LS state of the P450, and it is also not uncommon to see an increase in the absorbance of the δ peak accompanying the shift in the Soret peak to a longer wavelength. Typical P450 type II inhibitors include molecules such as imidazole and cyanide.

In both the above cases, the spectral changes upon ligand/substrate binding provide a useful tool for characterizing the binding of different compounds to a P450. With optical titrations, the changes that occur can be used to determine dissociation constants (\( K_d \) values), defined as the ligand concentration required to occupy half the active site binding sites.
1.6.1. The prevalence of P450s in nature

The P450s are among the oldest and most widespread genes, occurring in all kingdoms of life, from simple bacteria through to higher organisms such as Homo sapiens [94]. Because of this, it is thought that the P450 must have originated in an ancient, common ancestor. Today there are over 48,000 reported P450s (according to the SCOP website on 09.04.2013) [95]. A point of interest is the diversity of the numbers of P450s found in different organisms, which does not appear to correlate with the complexity of the organism. Humans contain 57 P450 genes [96], Arabidopsis thaliana contains 272 P450 genes [97], Drosophila melanogaster contains 87 P450 genes [98] and the pathogen Mycobacterium tuberculosis contains 20 P450 genes [99]. However P450s in bacteria are generally much lower in number than those found in eukaryotes, although the numbers of P450 genes in bacteria relative to genome size is often much greater in bacteria compared to eukaryotes [100]. The sheer volume of different P450 enzymes illustrates the wide variety of roles that these enzymes have in nature.

Most human P450s are found in the liver, where they perform a detoxification role for both physiological and xenobiotic compounds [101]. Typically this is achieved by the oxidation of hydrophobic compounds to make them more soluble, enabling them to be excreted in urine or bile [102]. The addition of an oxygen atom (often by hydroxylation) can allow the enzymatic conjugation of the P450 product to other soluble molecules (e.g. glutathione, sulfate, UDP-glucuronide) to further improve the efficiency of their removal from the body [103]. As there are a wide range of potential molecules to convert, a number of mammalian P450s exhibit broad (sometimes overlapping) substrate specificity. This allows for efficient metabolism and detoxification of a large range of molecules [104]. CYP1A1 is a good example of this “loose” substrate specificity, as it is able to metabolise more than 20 substrates [105]. Like many mammalian P450s, human P450s are difficult to study as a result of their lack of solubility, due to a lipid transmembrane anchor region at the N terminus of the protein. For some mammalian P450s, it has been possible to remove the N terminal anchor to improve protein solubility and allow for easier characterization [106, 107]. Making soluble mammalian P450s is beneficial and has led to an increased understanding of
how they bind to substrates and inhibitors, mainly through enabling their crystallisation and structural elucidation by P450 crystal X-ray diffraction.

Recently, an increasing number of studies have focused on investigating the roles of P450s in insects and plants. As already discussed, plants often contain large collections of P450 genes, which can represent about 1% of the plant protein coding genes [108]. Plant chemistry is perhaps much more complex than our own, which explains the reason for so many P450s. The P450s play key metabolic roles and have functions in enabling plants to defend themselves with physical and chemical deterrents (e.g. production of flavonoids [109] and lignin [110]), as well as having roles in cell growth and in signalling (e.g. CYPs 94C1 and 94B1 are involved in the biosynthesis of the signal transducer and plant hormone jasmonic acid in *A. thaliana* [111]). CYP71 is a family of P450s in plants that has continued to grow in number, and represents more than half of all the P450s in higher plants. In a broad sense, this group is involved in the metabolism of aromatic and aliphatic amino acid derivatives, alkaloids, fatty acids and hormones [108]. It is likely that, through gene duplications and diversification of duplicated genes such as the CYP171s, plants can adapt to new environments and exploit niche compounds, thus increasing the plants’ metabolic capabilities and competitiveness.

It is now thought that there are two broad groups of insect P450s. The first is involved in the metabolism of endogenous compounds (e.g. hormones, fatty acids and defensive compounds) and the second deals with external compounds (e.g. drugs, pollutants and insecticides) [112]. Many insects consume plants and seeds as part of their diets. Just as the plant uses P450s to produce defensive compounds, the insects use P450s to try and combat them. *Papillo pollyxenes* (the Eastern Black Swallow butterfly) is a good example of how insects can evade plant defenses. Plants produce furancocoumarins which have pro-oxidant activity that can damage the insect’s DNA. The *P. pollyxenes* produces CYP 6B1 and CYP 6B3 upon ingestion of furancocoumarin. These enzymes break down the plant compound via an epoxidation reaction of the furan ring, thus protecting the insect [112, 113]. The constant “war” between plants and insects is considered to be a key factor in driving the diversification of P450 activities, in order that one species can gain an advantage over the other and thrive in the environment.
1.6.2. Bacterial P450s
The most notable and well studied bacterial P450s are CYP101A1 \( (P450_{\text{CAM}}) \) from *Pseudomonas putida* and CYP102A1 \( (P450_{\text{BM3}}) \) from *Bacillus megaterium*, which hydroxylate camphor and fatty acids, respectively. Unlike their mammalian counterparts, bacterial P450s are typically highly soluble and contain no lipid anchor regions. Because of this, the majority of what we understand in terms of P450 structure originates from bacterial P450s (Section 1.7). Compared to mammalian P450s, in many cases the natural substrate for the bacterial P450s is not known. A good example is the bacterial P450 XplA \( (CYP177A1) \). XplA is currently being trialled as an enzymatic system for degrading explosives in polluted soils [114]. XplA binds the explosive RDX (Royal Demolition eXplosive) and degrades it by reduction of its nitro groups. However its “natural” substrate is not presently known. Bacterial P450s are often much more selective than hepatic P450s in their substrate range [115]. Because of their soluble nature, much effort has been put into engineering bacterial P450s to perform biotechnologically significant reactions, including the generation of metabolites of human drugs for use in diagnostics and pharmaceutical safety testing.

1.7. P450 Structure
In 1985 the full structure of perhaps the best known P450, CYP101A1, was obtained [116]. A heme domain structure for P450\(_{\text{BM3}}\) (CYP102A1) from *B. megaterium* was also solved in 1993 [117]. Subsequently many more P450 crystal structures have been reported, both with and without ligands/substrates bound (currently there are 27 unique P450 structures reported [118]). There are still many more bacterial P450 structures known compared to eukaryotic P450 structures, due mainly to the relative ease of crystallising these naturally soluble proteins. However, through increased understanding of the properties of eukaryotic P450s and technical developments (including improved engineering to remove the membrane anchor from mammalian P450s), the structures of several eukaryotic P450s have now been determined [119].

Although the amino acid sequences of the crystallised P450 proteins have a high level of variability, typically less that 20 % conservation between non family members, the
overall fold and shape of all the solved P450 structures are highly similar (Figure 1.10) [120]. P450s are globular proteins and generally have a triangular prism shape with a predominantly α helical structure with a smaller β sheet contribution. The α helices are labelled from the N terminus as A to L [121-123]. There is a large variation in the size of these helices, where those on the periphery tend to be smaller (e.g. A, B and H). The P450 contains a prosthetic heme group which is sandwiched between the I and L helices [124]. The heme is the catalytic centre for the P450 and above it sits the enzyme’s active site. The shape and chemical character of this active site plays a key role in determining the preferences for substrate binding. The topology is influenced by the side chains of the amino acids surrounding the active site which protrude into the pocket. The majority of the side chains are hydrophobic in nature, which explains why P450 substrates are typically hydrophobic and the active site is buried in the centre of the enzyme, away from the solvent [125].

Figure 1.10. Diagrammatic representation of the overall fold of P450\textsubscript{CAM} (CYP101A1). The major secondary structure elements are shown as blue cylinders (α helices) and brown arrows (β sheets). The heme cofactor is shown in red, and green strings represent interconnecting loops in the protein structure. The α helices are labelled with the conventional P450 nomenclature [122]. Image acquired from Munro \textit{et al.} (2006).
A vital part of the structure of any P450 is the heme binding motif that is conserved in the beta bulge region, which is adjacent to the L helix [123]. There is a sequence of amino acids which are conserved in virtually all known P450s (Figure 1.11). The heme binding motif is made up of about 10 amino acids, FxxGxxxCxG, where there is a stronger retention of some amino acids over others (the less conserved residues are denoted with an x). The heme iron is bound to the sulphur residue from the cysteine at the fifth axial ligand position. This amino acid is completely conserved in all P450s and its mutation results in the loss of heme binding. Also highly conserved are two glycines and a phenylalanine amino acid. The conservation of the glycine, which appears after the cysteine, can be explained by the need to form a tight bend to feed into the L helix [126]. The phenylalanine may have a role in modulating the electronic properties of the heme via bonding interactions with the cysteine thiolate ligand [127].

<table>
<thead>
<tr>
<th>P450</th>
<th>Sequence</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP51</td>
<td>CYP121 TSHLAQGCGHFCPSALGR</td>
<td>352</td>
</tr>
<tr>
<td>CYP101A1</td>
<td>VSHTTQGHCLQHLAR</td>
<td>364</td>
</tr>
<tr>
<td>CYP102A1</td>
<td>HAFKFGCQRACIQQFAL</td>
<td>410</td>
</tr>
<tr>
<td>CYP172A1</td>
<td>DAYLPGVSCICQGFAM</td>
<td>405</td>
</tr>
</tbody>
</table>

**Figure 1.11. Heme binding motifs of selected bacterial P450s.** Multiple sequence alignments of the conserved amino acids involved in the binding of the heme prosthetic group from CYP51 (Candida albicans), CYP121 (M. tuberculosis), CYP101A1 (P450cam from P. putida), CYP102A1 (P450BM3 from B. megaterium) and CYP172A1 (C. jejuni) are shown. Highly conserved amino acids are highlighted, 2 glycines and 1 phenylalanine (green) and the implicitly conserved heme coordinating cysteine residue (yellow).

### 1.8. Catalytic cycle of P450s

P450s are classed as monooxygenases as they abstract a single oxygen atom, from molecular oxygen, and insert it (usually) into a C-H bound of the substrate. The general reaction catalysed is displayed below.

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

For the reaction, the enzyme requires a reducing agent, most commonly NAD(P)H. It delivers two electrons and one proton (as a hydride ion) to a redox partner for the P450s,
and this protein either communicates directly with the P450 or indirectly via a second redox partner as a conduit to the P450. The consecutive reduction of the resting (ferric) P450 by two electrons allows for the bind and activation of molecular oxygen and for the incorporation of a single oxygen atom into the substrate bound close to the heme centre.

Much of what we understand in the catalytic cycle of P450s comes from the isolation of the intermediates from the reaction. Much of the work has been achieved using the P450\textsubscript{CAM} enzyme, and a current view of the cycle is depicted below (Figure 1.12). Recent studies from Michael Green and co-workers provided compelling evidence for the identification of compound I – the transient ferryl-oxo intermediate that is considered to be the species that introduces oxygen into the substrate [128, 129]. Thus, all species indicated in Figure 1.12 have now been identified spectroscopically.

Under resting conditions the distal site (sixth ligand position) on the heme is typically occupied by a H\textsubscript{2}O molecule, with the oxygen atom coordinating the heme iron. The water is displaced when a substrate binds in the active site pocket, and this usually causes the heme iron to move from the LS to the HS state (i). This increases the redox potential of the heme iron and allows the transfer of the first electron, reducing the iron from Fe\textsuperscript{3+} to Fe\textsuperscript{2+}. This electron comes from either NADH or NADPH and is delivered by a redox partner (ii). The ferrous iron binds dioxygen to create an oxyferrous complex (iii). This complex is the last relatively stable intermediate of the cycle [120]. This complex is a good electron acceptor and a second electron is transferred from the redox partner, leading to the formation of a ferric peroxy intermediate (iv). Rapid protonation creates the hydroperoxy-ferric intermediate otherwise referred to as Compound 0 (v). Delivery of a second proton leads to the collapse of Compound 0 and the O-O bound undergoes heterolysis to leave a ferryl (IV) iron-oxo complex, also called compound I, and the release of a water molecule (vi). The iron-bound oxygen is highly reactive and is inserted into the substrate to create the oxidized product (vii). The product is then released from the active site and is replaced by a water molecule, leaving the enzyme in the resting state and primed for another round of catalysis (viii). Due to the generation of highly reactive intermediates the notion of a completely cyclic system is not always
true and unwanted side reactions can occur (dotted lines in figure 1.12). Firstly, autooxidation of the oxyferrous complex can occur, releasing the bound oxygen as a superoxide, and returning the P450 to a ferric-heme state. With the second protonation event there is the possibility to get a reverse “peroxide shunt” reaction [130]. This leads to the cleavage of the Fe-O bond and the release of H₂O₂, returning the P450 back to its ferric-heme state. Finally the cycle can be broken via the oxidase shunt. This occurs if the oxygen atom of compound I is reduced to form water and the P450 is again returned to the ferric-heme state.

Although all the steps within the cycle are important in a traditional P450 reaction, the complete cycle does not always occur in special types of P450s. OleTJE from a *Jeotgalicoccus* bacterial species is one such special case, which is able to utilize H₂O₂ as a source for electrons, oxygen and hydrogen to produce terminal olefins from long chain fatty acids, and thus does not require electron delivery from redox partners [131]. This P450, and other peroxygenases, utilize the “peroxide shunt” pathway naturally, whereas other P450s find this route inefficient and also destructive to the heme cofactor.
Figure 1.12. Diagrammatic representation of the P450 catalytic cycle. The cycle begins with the binding of substrate to the oxidized protein, which displaces the coordinated water molecule (i). This allows reduction with the first electron enabling the binding of O₂ (ii, iii). The ferrous dioxygen species formed accepts another electron (iv) and two protons (v, vi). The loss of water leads to the formation of compound I (vi). The insertion of an oxygen atom into the substrate generates the product and returns the enzyme back to its resting state with water once again ligating to the heme iron (vii, viii). The reaction paths shown with dotted lines illustrate how the cycle can be disrupted with loss of superoxide, peroxide or water. However, in the case of the “peroxide shunt”, the reaction cycle can be driven in the productive direction with the addition of peroxide at the stage ii. Image adapted from Munro et al. (2007).
One very important step in P450 catalysis is the insertion of the oxygen into the substrate by compound I. Studies have indicated that this proceeds via a hydrogen abstraction/oxygen rebound mechanism (Figure 1.13) [132]. The high valent compound I is able to abstract the hydrogen from the substrate to form a ferryl hydroxyl species called compound II and a substrate radical. This radical then attacks the ferryl hydroxyl species to create the product and restores the iron to a ferric, resting state.

![Proposed mechanism for hydrogen extraction during substrate catalysis by P450s](image)

**Figure 1.13. Proposed mechanism for hydrogen extraction during substrate catalysis by P450s.** From left to right: Compound I extracts the hydrogen from the substrate to create a substrate radical and a protonated ferryl-oxo intermediate (centre, referred to as compound II [133]). The substrate radical then attacks the intermediate to produce a hydroxylated product and restores the heme to its resting state.

The catalytic cycle is a complex and energy demanding process, and so it is important that the cycle is well regulated. Control is often achieved based upon redox potentials. The Fe$^{3+}$ iron has a relatively negative potential in the substrate-free state. For example, P450 BM3 has a resting ferric heme iron midpoint potential of -427 mV [134]. The reducing power of NAD(P)H is approximately -320 mV. This means that it is difficult for electrons to flow from the reducing partner to the P450 in this state. It has also been reported that solvent exposure of the active site can alter the heme iron potential [135]. Under resting conditions the P450 has a high solvent exposure due to the water molecules in the active site. However, a number of these water molecules are displaced as the substrate binds. The sixth axial ligand bound water is one of these displaced waters, and the formation of the pentacoordinated species leads to a spin state change towards HS in the ferric P450. The change in spin state and loss of axial ligand results in the iron moving out of the plane of the porphyrin ring, and as electrons redistribute into...
the $e_g$ 3d orbital this increases the iron’s ionic radius [136]. This results in the alteration of the midpoint potential of the heme iron (to a much more positive potential), which in turn facilitates the transfer of electrons from NAD(P)H via one or more redox partner proteins.

The change in spin state is explained by ligand field theory. Whilst the P450 is in its resting state, the heme is in an octagonal geometry, with four ligands from the porphyrin ring and two axial ligands from the sulphur of the cysteine and from the oxygen atom of the water. The loss of the water following substrate binding results in the heme iron moving from LS to a HS state. This involves a reorganization of the heme iron 3d orbital electrons. In the resting state, LS, the d orbital electrons are in the ferric $S=1/2$ configuration, and in HS state the electrons are in the $S=5/2$ configuration (Figure 1.14).

Figure 1.14. The P450 heme iron 3d orbital arrangement for the low spin and high spin ferric forms. In the low spin state all the electrons occupy the lowest energy level as $\Delta_{\text{Oct}}$ (the difference in energy between $t_{2g}$ and $e_g$ orbitals) is large. For the high spin state the electrons are distributed between the $t_{2g}$ and $e_g$ orbitals, as $\Delta_{\text{Oct}}$ is decreased.

The important factor in determining whether the ferric heme iron is in the LS or HS state is the energy difference ($\Delta_{\text{Oct}}$) between the d orbital levels $t_{2g}$ and $e_g$. If $\Delta_{\text{Oct}}$ is large then more energy is required to promote electrons into the $e_g$ orbitals. The presence of substrates and ligands to the heme iron affects the size of $\Delta_{\text{Oct}}$. Some ligands exert strong fields, including CN$^-$ and imidazole (which ligate the ferric heme iron), and increase the energy difference between the two sets of orbitals, which reinforces the LS state. Weaker field ligands will reduce the energy difference, making it more likely for the heme to move towards a HS state. The dissociation of the water ligand to the heme
iron also decreases the energy difference and so its removal aids the electronic transition towards the HS state.

**1.9. Redox partners**

Virtually all P450s require two protons and two electrons to oxidise substrates into products. The electrons come almost exclusively from NAD(P)H. P450s cannot abstract electrons directly from the NAD(P)H. For this they require NAD(P)H-binding redox partners. For many years it was considered that there were only two classes of redox partners, and although the numbers of groups has increased, these two initial classes still represent the vast majority of P450 redox systems.

Class I systems include the majority of all bacterial P450 systems, along with mitochondrial P450 systems. Class I systems are made up of three components, a FAD-containing reductase, a ferredoxin (often possessing a 2Fe-2S cluster, although proteins containing 3Fe-4S and 4Fe-4S clusters have been reported [137]), and finally the P450 itself. The electrons are transferred from NAD(P)H, by hydride ion transfer, onto the FAD cofactor of the reductase, which results in a 2 electron reduction of the flavin quinone to a hydroquinone form. The electrons are then passed, one at a time, to the ferredoxin (which is a single electron carrier), which can then reduce the P450 in two single electron steps. In bacteria all the proteins in the pathway are soluble, but in eukaryotes only the ferredoxin protein is soluble, and the other two components contain membrane anchor regions.

Class II systems are the most common form of P450 systems in eukaryotes. These are two component systems with a FAD- and FMN-containing reductase (cytochrome P450 reductase or CPR) and the P450. These two components are both membrane bound in eukaryotes. The CPR is able to directly transfer the reducing equivalents to the P450. As with class I systems it is the FAD that accepts the electrons from NAD(P)H and then passes them onto the FMN (again, one at a time). It is believed that the CPR arose from a fusion of two proteins – due to the similarity of the domains to ferredoxin reductase and flavodoxin proteins [138]. To date only one prokaryotic class II system with
separate reductase and P450 components has been indentified, in *Streptomyces carbophilus* [139]. Unlike the eukaryotic systems, the bacterial components were both found to be soluble proteins.

In 1981, a discovery by Richard Ruettinger and Armald Fulco would change our understanding of the variability that can exist among redox systems [140]. From *Bacillus megaterium* they isolated a P450 fused to a CPR by a short linker, and further analyses revealed that neither domain contained any membrane anchor regions [141]. This enzyme was called P450$_{\text{BM3}}$ (CYP102A1) and is now known to have the fastest substrate oxidation rate reported for a P450, with a turnover rate of over 17,000 min$^{-1}$ for arachidonic acid [142]. It is able to achieve this fast rate as electron supply, the main rate limiting step in most P450 enzymes, is rapid in BM3. The fusion of the redox partner to the P450 and the efficient electron transport through the CPR module means that the flavin to heme electron transfer rate is much faster than in eukaryotic P450 enzymes. As depicted in the catalytic cycle, the delivery of electrons and protons at the correct time is important for the reaction to proceed efficiently. This is carefully controlled in BM3 and there is a strong coupling of NADPH oxidation to fatty acid substrate hydroxylation.

Since the discovery of this first “outlier”, many more novel redox system classes have been discovered, and at least ten such systems have been characterized. Table 1.3 highlights that the early classification system is too simplistic and outdated [94]. The increase in the numbers of classes is mainly due to the discovery of fusion systems (e.g. V –VIII in the table). Unlike P450$_{\text{BM3}}$ many of these new classes are not completely self-sufficient, and some require new and unexpected types of redox partners. The CYP116B family (VII) is completely self-sufficient and requires only substrate, oxygen and NAD(P)H for intramolecular electron transfer activity.
Chapter 1. Introduction

Class | Electron transport chain | Example organism/Enzyme
--- | --- | ---
I | NAD(P)H→[FdR]→[Fdx]→[P450] | *Pseudomonas putida, CYP101A1 [94]*
II | NADH→[CPR]→[P450] | *Streptomyces carbohylus, CYP105A3 [139]*
III | NAD(P)H→[FdR]→[Fldx]→[P450] | *Bacillus subtilis, CYP107H [143]*
IV | Pyruvate, CoA→[OFOR]→[Fdx]→[P450] | *Sulfolobus solfataricus, CYP119 [144]*
V | NADH→[FdR]→[Fdx–P450] | *Methylococcus capsulatus, CYP51 [145]*
VI | NAD(P)H→[FdR]→[Fdx–P450] | *Rhodococcus rhodochrous, XplA [146]*
VII | NADH→[PFOR–P450] | *Rhodococcus strain NCIMB, CYP116B2 [147]*
VIII | NADPH→[CPR–P450] | *Bacillus megaterium, CYP102A1 [94]*
X | [P450] | *Arabidopsis thaliana, CYP74A [94]*

**Table 1.3. Cytochrome P450 redox systems.** The ten currently characterised P450 redox systems are shown, along with details of the components of their respective electron transport chains, and an example of an organism that contains the relevant type of P450 redox system. FdR, ferredoxin reductase; Fdx, ferredoxin; CPR, cytochrome P450 reductase; Fldx, flavodoxin; OFOR, 2-oxoid ferredoxin oxidoreductase; PFOR, phthalate family oxygenase reductase.

### 1.10. Flavin-containing proteins

Flavins are found in all kingdoms of life and have a number of roles, including electron transport. Flavins are composed of a 7,8-dimethylisalloxazine ring system, and include the molecule riboflavin, and its derivatives: flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) (Figure 1.15) [149]. These cofactors are predominantly non-covalently attached to their proteins, just as the case for the heme group of the P450 [150]. Flavins can exist in three different redox states, oxidized, semiquinone (one electron reduced) and hydroquinone (two electron reduced). Flavoproteins can exploit this ability to facilitate the flow of one or two electrons. The environment that the protein exerts onto the flavin affects the stability of the three redox states, which in turn alters its ability to accept and release electrons [151]. All characterised flavoproteins have a midpoint potential in the range between -440 mV to +130 mV vs. the standard hydrogen electrode (SHE) [150]. Studies have revealed that the binding of pyridine nucleotides (NAD(P)H) alters the redox potential of the flavin in reductase/oxidase enzymes that interact with these cofactors, which often allows electron transfer to occur more readily than would be predicted based on the potentials of the isolated flavoprotein and cofactor [152]. As mentioned, the majority of flavins are non-covalently linked to the protein backbone. However, there are some flavoprotein systems where covalent
attachment of the cofactors occurs. *Vibrio cholerae* contains two flavoproteins, RnfD and RnfG, that covalently attach their flavin (FMN) moieties to the protein at threonine-187 and threonine-175, respectively [153]. These two proteins are involved in a complex of membrane bound proteins involved in transferring electrons from NADH to ferredoxin. FAD covalent linkage has also been observed in many proteins, e.g. in succinate dehydrogenase, where the FAD is covalent bound to a histidine through the 8-methyl group of the isoalloxazine ring [154]. The reason and potential benefit for this linkage is not yet understood, but it reasonable to suggest that covalent linkage may have a role in tuning the flavin redox potential for a specific redox function and/or increase the retention of the flavin cofactor, thus ensuring the protein remains functional.

Figure 1.15. Chemical structures of riboflavin and its derivatives FMN and FAD. Riboflavin is the simplest molecule of the group and consists of an isoalloxazine ring with a reduced ribose sugar linked at the N10 position. A single phosphorylation of the ribose creates FMN, which can be converted into FAD with the addition of a second phosphate group and nicotinamide base.
1.10.1. Ferredoxin/Flavodoxin reductases

Ferredoxin reductases (FNR) are proteins which contain FAD cofactors, and that mediate the delivery of electrons from NADPH to ferredoxin or flavodoxin partners. FNRs contain two domains, a N-terminal FAD binding domain and a C-terminal NADP+ binding domain, whose structure is well conserved in many dehydrogenases (Figure 1.16) [155]. The C-terminal domain contains the conserved Rossmann fold which is required for binding of nucleotides [156]. This fold consists of six parallel β sheets separated by α helices to create two βαβαβα motifs. The FAD is non-covalently bound to the reductase and its isoalloxazine ring is sandwiched between conserved aromatic residues. A commonly conserved motif is the RxYS(T) (brackets indicate variation in the residue). In the crystal structure of FNR from *E. coli* the tyrosine-52 side chain positions itself parallel to the isoalloxazine ring of the FAD. The conserved arginine-50 forms a hydrogen bonds to the phosphate groups on the FAD and the conserved serine-53 forms a hydrogen bond to the isoalloxazine ring (at the nitrogen at position 5 on the ring) [157]. In rat CPR the mutation of the corresponding arginine residue to a glutamate resulted in a 25,000 fold decrease in FAD affinity [158]. A second aromatic residue then occupies the other face of the FAD isoalloxazine ring to ensure it is kept in place and to regulate its redox properties. Mutations of either aromatic amino acid also reduced the affinity of FAD binding [158].

There are two other conserved motifs, GxxT/S and GxGxxP, where the former is involved in binding the pyrophosphate of either FMN or FAD. The position of the glycine, relative to the T/S is thought to be important in determining preferential binding of FMN or FAD [159]. The GxGxxP motif is located in the NADP+ binding domain and the amino acids in the motif interact with the pyrophosphate in NAD(P)H [157].
Figure 1.16. Ribbon cartoon of the *E. coli* ferredoxin reductase structure. The N-terminal (FAD binding) domain is in blue, and the C-terminal (NADP⁺-binding) domain is in red. The domains sandwiching the FAD and NADP⁺ are depicted as secondary structural models. The structure is of *Anabaena* (sp. pcc 7119) ferredoxin reductase (PDB code; 1GJR) and was drawn in PyMol.

1.10.2. Flavodoxins

Flavodoxin redox activity arises from its tightly bound FMN cofactor [160]. Flavodoxins are thought to have been an evolutionary solution reached for organisms requiring electron transfer when located in low iron environments, in which they are unable to synthesise ferredoxin iron-sulphur clusters (Section 1.10.3). Flavodoxins are found only in bacteria and not in eukaryotic systems. However a descendant flavodoxin does appear in fusion proteins in some eukaryotes, including human cytochrome P450 reductase (CPR) [161]. Flavodoxins can be separated into two different groups, short chain flavodoxins and long chain flavodoxins, which differ by the presence of a 20 residue loop (of unknown function) in the latter class [160]. Removal of the loop from a long chain flavodoxin in *Anabaena* resulted in weaker binding of FMN to the protein, which the authors suggested was a result of conformational changes to the “loopless” protein, and proposed that the loop may be used for reductase partner recognition [162].
Contrary to FNRs, which can accept one or two electrons at once, flavodoxins typically only act as single electron donors or acceptors. In solution, the reduced form of free FMN is unstable and reoxidises rapidly [163]. In complex with a protein the one electron reduced (semiquinone) and two electron reduced (hydroquinone) forms can be stabilised, with the level of stabilisation of one or other form dependent on the binding environment. Mutational studies on a variety of bacterial flavodoxins highlighted that the conserved tyrosine residue, key in packing the isoalloxazine, is important for stabilising the different redox states of FMN. It also shifts the semiquinone/hydroquinone redox couple to much more negative redox potentials, compared to free FMN [164].

FMN incorporation is driven by the binding of the isoalloxazine ring to the protein. Mutations of key tryptophan and tyrosine residues (sandwiching the isoalloxazine ring) of a flavodoxin protein from *Anabaena* lowered the rate of complex formation (by 10 to 100 fold) [165]. Once the binding of the ring takes place, the incorporation of the FMN is strengthened by interactions of the phosphate and ribityl sections of the FMN to key amino acids within the protein [160]. Binding of the FMN helps to stabilize the flavodoxin protein and increases its thermostability by preventing the exposure of key residues [166].

1.10.3. Iron-sulphur containing proteins

Ferredoxins are important biological electron carriers, and contain bound iron-sulphur clusters [137]. Clusters involved in electron transport contain either rhomboidal [2Fe-2S] or cuboidal [3Fe-4S] or [4Fe-4S] structures (Figure 1.17). Ferredoxins can exist in two redox states (oxidised and reduced) and as such are involved in single electron transfer [137]. Studies on P450<sub>CAM</sub> and in mitochondrial systems show that the ferredoxin binds in a 1:1 manner to both the FAD-containing reductase (adrenodoxin/putidaredoxin reductase) and the P450 protein partners [167]. Fe-S clusters typically have low redox midpoint potentials, with [2Fe-2S] clusters in proteins having midpoint potentials ranging from -100 to -450 mV vs. SHE [168].
Figure 1.17. Different types of naturally occurring iron-sulphur clusters. Structural arrangements of [2Fe-2S], [3Fe-4S] and [4Fe-4S] clusters are shown.

1.11. Inhibitors of P450s

Azoles are drugs that were first identified for their antifungal capabilities. The first azole drug to be described was benzimidazole, in 1944 [169]. However, research into the potential of azoles did not attract major attention for several years. In the 1960s the well known (and still used) azoles; econazole, clotrimazole and miconazole were produced (Figure 1.18) [170]. Azoles target the fungal P450, CYP51, a lanosterol demethylase, which is responsible for the synthesis of the fungal sterol ergosterol. Ergosterol forms an integral part of the plasma membrane and is involved in maintaining membrane fluidity and stability (much like cholesterol in human membranes) [171]. The unpaired electrons on an exposed azole nitrogen group are able to bind directly to the heme iron, preventing the activation of oxygen which is required for the P450 to function. Recently it was reported that azole uptake by C. albicans occurs by facilitated diffusion and not just simply by passive diffusion [172]. This might occur if the azole structure closely matches the structure of a molecule that the fungus utilizes naturally.
Figure 1.18. Chemical structures of A) econazole and B) clotrimazole. The exposed N atom in the imidazole ring coordinates to the P450 heme iron to inactivate the enzyme.

1.12. Aims of the project

*C. jejuni* is a human pathogen that is one of the most commonly reported causes of gastroenteritis and has been identified by the WHO as an organism that requires further study in order to increase our knowledge and understanding of the mechanisms this bacterium employs to cause infection. Cytochrome P450s are less common in prokaryotes than eukaryotes, but have been shown, in some cases, to be essential for bacterial survival (e.g. CYP51 in fungi and CYP121 in *M. tuberculosis*) or are involved in biosynthetic pathways. As the P450 protein, CYP172A1, of *C. jejuni* has only been a recent discovery, there is still a wealth of knowledge to be extracted. Of particular interest is to identify the role that the protein plays and to ascertain why the gene is not present in all Campylobacter species.

There are two main objectives to the work in this thesis. The first involves examining the role of the novel protein *in vivo*. An insertional knockout mutant of the P450 gene (*cj1411c*) was constructed in order to prevent its expression by the host cell, so allowing for comparative studies between the wild type and the mutant to be performed. As the role of the P450 enzyme is unknown, general microbiological techniques including growth assays, morphological analysis and antimicrobial sensitivity assays were performed for both the wild type and mutant strains in two well characterised *C. jejuni* strains; 11168H and 81-176. Any phenotypes observed were further examined by the
reintroduction of the P450 gene to the mutant cell line by complementation. Due to the genetic location of the \textit{cj1411c} gene, a large body of the work was focussed on investigating the effects of P450 gene disruption on the bacterial cell surface, with particular attention directed at the capsule, lipooligosaccharide and flagellum. These cellular components were investigated using classical microbiological techniques, as well as more recent methods for bacterial characterisation and discrimination, focusing on novel spectroscopic techniques.

The other major part of this work focused on studying the CYP172A1 P450 enzyme \textit{in vitro}. By using cloning, expression and purification techniques, a recombinant CYP172A1 was produced in large quantities to allow for confirmation that the protein was indeed a P450 and to investigate its biophysical characteristics. This P450 was then investigated for its interactions with potential substrate and inhibitor type ligands, and to calculate $K_d$ values for these molecules in order to probe its active site structure and molecular preference. In particular, ligand binding studies of azole compounds, which are increasingly seen as potent P450 inhibitors, were performed. As the P450 contains low amino acid sequence similarity to other P450s of known structure, attempts were made to determine the CYP172A1 structure using X-ray crystallography. A CYP172A1 structure should reveal the P450 active site structure and provide insights into substrate and inhibitor selectivity. To further aid this goal, compound library screening studies were done in order to define type I and type II molecules binding to CYP172A1.

The work carried out serves as the first characterisation of the only P450 enzyme found in the human pathogen \textit{Campylobacter jejuni}. 
Chapter 2. Materials and Methods

2.1. Materials

All chemicals and reagents used for experimental procedures were purchased from, Bioline (London, UK), bioMérieux (Basingstoke, UK), ForMedium (Hunstanton, UK), Fisher Scientific (Loughborough, UK), Invitrogen (California, USA), Merck (New Jersey, USA), NEB (Ipswich, MA, USA), Oxoid (Basingstoke, UK), Qiagen (Manchester, UK) or Sigma-Aldrich (Poole, UK).

2.2. Growth media

All liquid media were sterilized at 121°C, at 15 psi for 20 minutes in an autoclave (LTE Scientific, Oldham, UK). Once the media had cooled to below 50 °C, antibiotic(s) were added where appropriate (Table 2.1) prior to inoculation of media with bacterial cells.

Luria-Bertani (LB) Medium (ForMedium)

For every litre of distilled water ($dH_2O$), the following were added:
- 10 g Bacto-tryptone
- 5 g Bacto-yeast extract
- 10 g NaCl

2xYT Medium (ForMedium)

Per litre of $dH_2O$, the following were added:
- 16 g Bacto-tryptone
- 10 g Bacto-yeast extract
- 5 g NaCl
Terrific Broth (ForMedium)
Per litre of $dH_2O$, the following were added:

- 11.8 g Peptone
- 23.6 g Yeast extract
- 9.4 g dipotassium hydrogen phosphate
- 2.2 g potassium dihydrogen phosphate

SOB Medium
Per litre of $dH_2O$, the following were added:

- 20 g Bacto-tryptone
- 5 g Bacto-yeast extract
- 0.5 g NaCl

Mueller Hinton (MH) Broth (Oxoid)
Per litre of $dH_2O$ the following were added:

- 17.5 g Acid Hydrolysate of Casein
- 3.0 g Beef Extract
- 1.5 g Starch

All agar based media were sterilized at 121 °C, at 15 psi for 20 minutes. Then the agar solution was cooled to below 50 °C, antibiotic(s) were added where appropriate (Table 2.1) and the mixture poured into standard 90 mm diameter plastic petri dishes (Sarstedt, Germany). Once the agar had set, plates were stored (4 °C) for up to four weeks until use.
LB Medium Agar Plates (ForMedium)

Per litre of $dH_2O$, the following were added:

- 10 g Bacto-tryptone
- 5 g Bacto-yeast extract
- 10 g NaCl
- 10 g Bacto-Agar

0.4 % and 1 % MH agar (Oxoid)

Per litre of $dH_2O$, the following were added

- 17.5 g Acid Hydrolysate of Casein
- 3.0 g Beef Extract
- 1.5 g Starch
- 4 g (0.4 %) or 10 g (1 %) Bacto-Agar.

Columbia Blood Agar (CBA) (Oxoid)

CBA was prepared with the addition of defibrinated horse blood (TCS Biosciences, Buckingham, UK) at 5 % v/v when the medium was below 50 °C.

Per litre of $dH_2O$, the following were added

- 23.0 g Special nutrient substrate
- 1.0 g Starch
- 5.0 g NaCl
- 13.0 g Agar
2.3. Antibiotics

Antibiotics used in the course of the study and their stock and working concentrations are given in Table 2.1

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Stock solution concentration and preparation</th>
<th>Final concentration (μg/ml)</th>
<th>Liquid culture</th>
<th>Agar plates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin (Amp)</td>
<td>50 mg/ml in dH₂O. Filter sterilized and stored at -20 °C</td>
<td>50</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Carbenicillin (Carb)</td>
<td>50 mg/ml in dH₂O. Filter sterilized and stored at -20 °C</td>
<td>50</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Chloramphenicol (Chl)</td>
<td>34 mg/ml in ethanol. Stored at -20 °C. Protected from light.</td>
<td>34</td>
<td>34-68</td>
<td></td>
</tr>
<tr>
<td>Kanamycin (Kan)</td>
<td>30 mg/ml in dH₂O. Filter sterilized and stored at -20 °C</td>
<td>30 µg/ml for cells containing kan&lt;sup&gt;R&lt;/sup&gt; plasmid. 30 µg/ml for chromosomal kan&lt;sup&gt;R&lt;/sup&gt; gene. 15 µg/ml for cells containing chromosomal kan&lt;sup&gt;R&lt;/sup&gt; gene.</td>
<td>60 µg/ml for kan&lt;sup&gt;R&lt;/sup&gt; gene on plasmid. 30 µg/ml for chromosomal kan&lt;sup&gt;R&lt;/sup&gt; gene.</td>
<td></td>
</tr>
<tr>
<td>Rifampicin (Rif)</td>
<td>10 mg/ml in 65% ethanol, 0.17 M NaOH. Stored at -20 °C and used within 2 weeks. Protected from light.</td>
<td>200</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Tetracycline (Tet)</td>
<td>5 mg/ml in 50% ethanol. Stored at -20 °C. Protected from light</td>
<td>34</td>
<td>68</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1. Antibiotics used in the course of this study. The concentrations of their stock and working concentrations for both liquid and solid media are given.
2.4. *Campylobacter jejuni* strains

*Campylobacter jejuni* strains 11168H and 81-176 cells, the genetically manipulated *cj1411c* insertional knockout and complement strains, and other *Campylobacter jejuni* strains used during the study are listed in Table 2.2. Strains were obtained from or created in the Linton laboratory during the course of this study.

<table>
<thead>
<tr>
<th>Strain/genotype</th>
<th>Description</th>
<th>Short hand identification</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. jejuni</em> NCTC 11168H</td>
<td>Wt strain</td>
<td>PE1</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168H Δcj1411c::kan</td>
<td>KO strain</td>
<td>PE2</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168H Δcj1411c::kan Δcj0223::chl/cj1411c</td>
<td>Complement strain</td>
<td>PE3</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 ΔFlaA::kan</td>
<td>FlaA mutant</td>
<td>ΔFlaA</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 ΔkpsM::kan</td>
<td>kpsM mutant</td>
<td>ΔkpsM</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 Δ1139::kan</td>
<td>1139 mutant</td>
<td>Δ1139</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 P1</td>
<td>PS 1</td>
<td>P1</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 P1 ΔkpsM::kan</td>
<td>PS 1 kpsM mutant</td>
<td>P1a4</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 P2</td>
<td>PS 2</td>
<td>P2</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 P2 ΔkpsM::kan</td>
<td>PS 2 kpsM mutant</td>
<td>P2a4</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 P10</td>
<td>PS 10</td>
<td>P10</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 P10 ΔkpsM::kan</td>
<td>PS 10 kpsM</td>
<td>P10a4</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 G1</td>
<td>GB 1</td>
<td>G1</td>
</tr>
<tr>
<td><em>C. jejuni</em> NCTC 11168 G1 ΔkpsM::kan</td>
<td>GB 1 kpsM mutant</td>
<td>G1a4</td>
</tr>
<tr>
<td><em>C. jejuni</em> 81-176</td>
<td>Wt strain</td>
<td>PE4</td>
</tr>
<tr>
<td><em>C. jejuni</em> 81-176 Δcj1411c::kan</td>
<td>KO strain</td>
<td>PE5</td>
</tr>
<tr>
<td><em>C. jejuni</em> 81-176 Δcj1411c::kan Δcj0223::chl/cj1411c</td>
<td>Complement strain</td>
<td>PE6</td>
</tr>
</tbody>
</table>

Table 2.2. List of *Campylobacter jejuni* NCTC 11168H and 81-176 strains and mutants used in this study. The wild type (Wt), Δcj1411c::kan (KO) and Δcj1411c::kan Δcj0223::chl/cj1411c (Complement) forms of *C. jejuni* NCTC 11168H and 81-176 strains were used in these studies, with the latter two mutants created during the duration of this study. Each strain was given a short hand identification tag. Wt, wild type; KO, knockout; PS, Penner serotype; GB, Guillain-Barré

Other *Campylobacter* species used were as follows; *C. coli* RM2228, *C. lari* RM2100 (Wren laboratory, London School of Hygiene and Tropical Medicine, UK), *C. fetus* subspecies *fetus* NCTC 10842 and *C. upsaliensis* NCTC 11541(Dr Andy Lawson, Health Protection Agency, UK).

All of the above *Campylobacter* species are fast growing and do not require additional hydrogen as part of their growth conditions.
*Campylobacter* species, unless otherwise stated, were grown on CBA medium supplemented with 5% (v/v) horse blood at 42 °C for 48 h under microaerobic conditions (5% CO₂, 5% O₂, 90% N₂) in a MACS VA500 workstation (Don Whitley Scientific, West Yorkshire UK).

### 2.5. Purification of *C. jejuni* genomic DNA

DNA was extracted using a Puregene® extraction kit under the manufacturer’s protocol (Gentra, Minnesota, USA).

### 2.6. Amplification of target DNA by PCR

Oligonucleotide primers designed to bind specifically to target DNA sequences (see Table S1) were used to amplify DNA of interest. Template DNA (100 ng) was added to a PCR master mix (1x NH₄ buffer, 1.5 mM MgCl₂, 0.8 mM dNTPs, 2.5 U Taq DNA polymerase (all from Bioline, UK)) along with the required primers (20 pM) to a total volume of 50 µl in 200 µl PCR tubes. These tubes were placed in a PxE 0.2 thermal cycler (Thermo Scientific) and DNA amplified using a defined program (Table 2.3). For increased fidelity, 0.2 µl of Pwo DNA polymerase (10 U/µl) (Roche, Indianapolis, USA) was added, along with the Taq polymerase in selected PCR reactions. The annealing temperature (X) was set 5 °C below the lowest primer annealing temperature. The elongation time (Y) was dependent on the length of DNA to be amplified and typically one minute was sufficient for ~1000 nucleotides to be replicated.
Table 2.3. PCR program used to amplify target DNA. Amplification of DNA was achieved using a three step method with times and temperatures listed. X, the annealing temperature was set at 5 °C below the lowest $T_m$ of primers used. Y, the elongation time was dependent on the length of DNA to be amplified, typically one minute/1000 nucleotides.

### 2.7. Preparation of electrocompetent *C. jejuni*

*C. jejuni* cultures were grown for 24 hours on CBA, cells were then harvested and suspended in 1 ml of ice cold sucrose (272 mM) containing 15 % (v/v) glycerol. Cells were washed three times in this solution, with centrifugation at 10,000 g prior to resuspension. These electrocompetent cells were aliquoted into 50 µl batches and either used immediately or stored at -80 °C for future use.

Plasmid constructs (10 ng) were mixed with 50 µl of electrocompetent cells in 1 cm electroporation cuvettes (Bio-Rad, California, USA) and electroporated using a Bio-Rad Gene pulser II coupled to a Pulse controller II and Capacitance Extender plus (Bio-Rad). Settings were 2.5 kV at a resistance of 200 Ω and capacitance of 25 µF. Cuvettes were flushed with 200 µl SOB medium and grown on non-selective CBA for 16 hours. Cells were then plated onto appropriate antibiotic selective agar, and plates incubated under microaerobic conditions until colonies formed. Individual colonies were subcultured, the genotype verified by PCR, and bacteria stored as glycerol stocks in a -80 °C freezer by suspending cells in 30 % glycerol (v/v) in MH.
2.8. Sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis

For studies in Chapter 3, protein and cell samples were denatured in 1x SDS sample buffer (63 mM Tris HCl, 10 % glycerol, 2 % SDS and 0.0025 % Bromophenol Blue [173]) containing a 1 % (v/v) beta-mercaptoethanol solution, by heating to 95 °C for five minutes. Samples were briefly centrifuged in a bench top centrifuge and the supernatants were loaded onto a precast 10 or 12 % SDS-PAGE gel along with broad range pre-stained protein markers (Bio-Rad, California, USA). Electrophoresis was performed in an XCell Surelock™ Mini-Cell with 1x SDS-PAGE running buffer (25 mM Tris, 192 mM glycine and 0.1 % SDS). Electrophoresis was performed for one hour at 180 V and gels were stained using Instant Blue (Expedeon, Cambridge, UK.).

For the studies described in Chapter 4, protein and cell samples were denatured in 3x SDS sample buffer containing a Dithiothreitol (DTT) solution (125 mM) by heating to 95 °C for five minutes. Gels were run using precast 10 or 12 % SDS-PAGE gels (GeBa, Belgium and Bio-Rad, USA) and run in 1x SDS running buffer at 180 V for one hour. Broad range pre-stained protein markers (Bio-Rad, California USA) were also loaded onto the gel. After the allotted time required for resolution of proteins, the GeBa gels were stained in Coomassie Blue stain (0.1 % Brilliant blue G, 40 % ethanol (EtOH), 10 % acetic acid and 50 % H2O) for 20 minutes and destained using dH2O. Bio-Rad gels were visualized, without staining, on a Bio-Rad Gel Doc EZ Imager (Bio-Rad, USA). This was done to estimate protein purity as well as to identity specific protein bands.

2.9. Western blotting

In studies described in Chapter 3, SDS-PAGE gels were not stained with Instant Blue but were transferred onto Protran nitrocellulose membrane (Whatman, Maidstone, UK) using an XCell II™ Blot module (Invitrogen, California, USA) at 30 mV for one hour in 1x transfer buffer (25 mM Tris, 192 mM glycine in 20 % methanol (MeOH)). Membranes were blocked overnight (O/N) in phosphate buffered saline (PBS) (Sigma-Aldrich Ltd, Poole, UK) containing 5 % (w/v) milk solution (Premier International Food, Republic of Ireland). To detect proteins of interest, the blot was incubated with an
anti CYP172A1 antiserum (generated as part of this study, see Section 2.30) for one hour. The antiserum was diluted 1:10 with a TNT solution (25 mM Tris at pH 7.5, 160 mM NaCl and 0.5 % Tween-20) and 5 % milk. C. jejuni PE2 (at OD$_{600}$ = 1) suspended in PBS was lysed by freeze/thaw action using dry ice, and 500 μl of lysed cells were added per ml of antiserum/TNT solution. This solution was mixed for four hours at room temperature prior to use on the membrane, at 1:100 dilution in TPBS (Tween-20 at 0.01 % in PBS) with 3 % (w/v) milk. The antibody solution was removed and the membrane washed three times with TPBS. The secondary antibody, a fluorescently labeled anti-rabbit IgG (LI-COR, Lincoln, USA), was diluted 1:10,000 in TPBS with 3 % milk and incubated with the membrane for an hour. Membranes were washed three times with TPBS, once in PBS and immunoreactive proteins visualised using a LI-COR detector (LI-COR, USA) and the 800 nm channel detector.

In the studies described in Chapter 4, electrophoretically separated proteins in SDS-PAGE gels were not stained with Coomassie blue, but instead were transferred onto a PVDF (polyvinyl difluoride) membrane using a Trans Blot Turbo instrument (Bio-Rad, USA) set at 25 V and 1.3 A for seven minutes using a Bio-Rad quick membrane kit. The WesternBreeze$^\text{®}$ chemiluminescent kit and protocol (Invitrogen, USA) was used to detect antibody binding to target proteins. Following protein transfer, the membrane was placed into a 50 ml Falcon tube (Fisher Scientific, Loughborough, UK) and washed in 10 ml of blocker solution (5ml H$_2$O, 2 ml diluent a (concentrated saline solution with detergent), and 3 ml diluent b (concentrated Hammersten casein solution)) at 4 °C with gentle agitation for one hour. The blocker solution was decanted, the membrane briefly washed with deionised distilled water (ddH$_2$O) and exposed to the primary antibody at a 1:5000 dilution (in blocker buffer) for one hour at 4 °C. The primary antibody solution was poured off and the membrane washed four times with the supplied 1x antibody wash solution (saline solution with detergent) for five minutes. The membrane was immersed in an anti-mouse or anti-rabbit (where appropriate) antibody solution (alkaline phosphatase-conjugated, anti-species IgG (Invitrogen, USA)) for one hour. This solution was removed and the membrane washed three times for 20 minutes using the 1x antibody wash solution followed by a two minute wash with ddH$_2$O. Finally 2 ml of
alkaline phosphatase substrate (0.25 mM) (Invitrogen, USA) was incubated on the membrane for 2 minutes. Excess reagent was removed by blotting with filter paper and the membrane sealed in a plastic sheet and kept in the dark until it was developed. Under red light, an X-ray film (Kodak) was placed over the membrane for between 30 seconds and five minutes and the film developed using a X-ray film processor (AFP ImageWorks, New York, USA).

2.10. C. jejuni growth assays
Aliquots (100 µl) of C. jejuni strains and mutants of 11168H and 81-176 were inoculated into 21 ml MH broth and grown under microaerobic conditions for 24-48 hours with agitation (80 r.p.m) with use of appropriate antibiotics. After an initial growth period the optical density (OD$_{600}$) was measured and the culture used to inoculate a fresh flask of MH broth to an OD$_{600}$ of 0.05. Cultures were grown as described above, but without any antibiotics, and periodically OD$_{600}$ and cell viability was determined using the Miles and Misra technique [174]. In brief, cell samples were diluted ten-fold in MH and 3x 10 µl of cell suspensions were deposited onto CBA plates and incubated at 42 °C for two to four days prior to cell counting.

2.11. C. jejuni autoagglutination assays
Autoagglutination methodology was based on previous work [54, 175]. Briefly, C. jejuni 11168H and 81-176 strains were grown on CBA plates (with appropriate antibiotics where required) at 42 °C for 48 hours under microaerobic conditions. Cells were then resuspended in 25 ml of sterile PBS to an OD$_{600}$ of 1. The suspensions were distributed evenly into three 15 ml Falcon tubes (Fisher Scientific), and positioned vertically and stationary. The top 1 ml of suspension was removed periodically and the OD$_{600}$ recorded. The data was analysed in Microsoft Excel software and statistical analysis performed using the Student’s paired t-test (p≤0.05).
2.12. Electron Microscopy (EM)

*C. jejuni* strains were grown for 24 hours on CBA plates under microaerobic conditions at 42 °C. Using a small plastic loop bacteria were carefully removed from the plate surface and suspended in *ddH₂O* to an OD₆₀₀ of 0.1. Copper EM grids were coated in 0.3 % Formvar (w/v in chloroform) and then carbon coated to produce a surface for the cells to adhere to. Shortly before use, the grids were treated with glow discharge (set at 100 mV for 30 seconds under vacuum) to increase the hydrophilicity of the surface [176]. The grids were placed onto droplets of bacteria, washed three times with *ddH₂O* and then negatively stained with 1 % (w/v) uranyl acetate solution for 45 seconds. Excess stain was removed by blotting with filter paper and the grids dried for 24 hours. EM images were acquired on a FEI Tecnai 12 Biotwin Transmission Electron Microscope (FEI, Oregon, USA).

2.13. *C. jejuni* motility Assay

Motility tests were performed as previously described [177]. In brief, *C. jejuni* 11168H and 81-176 strains were grown as described previously. A 1 ml of cell suspension was centrifuged and resuspended in MH to an OD₆₀₀ of 0.1. From this cell suspension, 10 µl was point inoculated into the base of a 0.4 % MH agar plate with a 0.1-2.0 µl pipette tip. Plates were incubated for 48 hours at 42 °C under microaerobic conditions. Bacteria motility was measured as the diameter of the circular motile zone.

2.14. Minimum Inhibitory Concentration (MIC) assays for *C. jejuni*

MIC values were determined using the broth dilution method [178]. Firstly 100 µl serial dilutions of a range of antimicrobials (Table 2.4) were pipetted into individual wells in a 96 well microtitre plate (Fisher Scientific). Bacterial suspensions were prepared at 10⁵ cfu ml⁻¹ in MH (prepared from mid log growth liquid cultures) and 100 µl added to the individual wells. Control wells containing no antimicrobials and no cells were also included. The microtitre plates were incubated at 42 °C for 24 h in a microaerobic atmosphere using an anaerobic gas generating kit (Oxoid, UK). Cell growth within the 96 wells was measured on an automated plate reader (BioTek, Winooski Vermont,
USA) linked up to Gen5 computer software. MIC values were then determined by comparisons of cell growth in control and assay cultures.

<table>
<thead>
<tr>
<th>Antimicrobial</th>
<th>Working concentration range (μg/ml)</th>
<th>Mode of action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin</td>
<td>1024 - 2</td>
<td>Inhibits peptidoglycan synthesis.</td>
</tr>
<tr>
<td>Cefotaxime</td>
<td>256 - 0.5</td>
<td>Inhibits peptidoglycan synthesis.</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>8 - 0.015</td>
<td>Inhibits DNA replication.</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>32 - 0.625</td>
<td>Inhibits protein synthesis by binding to the 50S ribosomal subunit.</td>
</tr>
<tr>
<td>Gentamycin</td>
<td>16 - 0.031</td>
<td>Inhibits protein synthesis by binding to the 30S ribosomal subunit.</td>
</tr>
<tr>
<td>Nalidixic acid</td>
<td>1024 - 2</td>
<td>Inhibits DNA replication.</td>
</tr>
<tr>
<td>Polymyxin B</td>
<td>256 - 0.5</td>
<td>Alters ionic homeostasis and permeability at the cytoplasmic membrane.</td>
</tr>
<tr>
<td>Sodium deoxycholate</td>
<td>512 - 1</td>
<td>Lyses target cells by altering plasma membrane solubility.</td>
</tr>
<tr>
<td>SDS</td>
<td>25,000 - 1</td>
<td>Lyses target cells by altering plasma membrane solubility.</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>64 - 0.125</td>
<td>Inhibits protein synthesis by binding to the 30S ribosomal subunit.</td>
</tr>
</tbody>
</table>

Table 2.4. List of antimicrobials tested against *C. jejuni*. The working concentration was produced via serial dilution. The cellular targets for the antimicrobials are also described.

2.15. Polymyxin B susceptibility assays

1) E-test strips.

*C. jejuni* strains were in grown to mid log stage in MH under microaerobic conditions. Cells were adjusted to $10^5$ cfu ml$^{-1}$ in MH and 200 μl applied to a dry CBA plate. The cells were spread across the plate using a sterile glass spreader and were allowed to dry for 5 minutes. Using forceps, a single Polymyxin B E-test strip (bioMérieux, France) was carefully placed onto the agar surface and plates incubated under microaerobic conditions for 48 hours. MIC values were measured according to the manufacturer’s protocol.
2) Polymyxin B agar dilution
To molten CBA agar, a range of polymyxin B concentrations (8-16 µg ml\(^{-1}\)) were added prior to the agar setting. \textit{C. jejuni} strains were grown to mid log stage in MH cultures under microaerobic conditions. Cells were adjusted to 10\(^6\) cfu ml\(^{-1}\) in MH and sequentially diluted ten-fold three times and 10 µl of dilutions were spotted onto the polymyxin B-containing CBA plates and incubated for 48 hours. Cell viability was calculated using the Miles and Misra technique \cite{Miles1938}. The percentage growth inhibition at the different concentrations of polymyxin B was calculated using controls.

2.16. Disc diffusion assay for \textit{C. jejuni}
The disc diffusion assay was carried out a previously reported, to the National Committee for Clinical Laboratory Standards \cite{NCCLS1998}. Blank assay discs (Oxoid, USA) were soaked in 15 µl of the selected compound (typically 50 mg/ml and dissolved in DMSO), and were dried for a brief period. \textit{C. jejuni} was grown in MH until mid log phase and 1 ml of culture was removed and centrifuged in a bench top centrifuge at 3400 g for two minutes. The pellet was suspended in MH to an OD\(_{600}\) of 1 and 50 µl inoculated into 5 ml of molten 0.4 % MH agar at 37 °C. This suspension was poured onto the surface of a preset 1 % MH agar plate and allowed to set. Once dry, using forceps, the assay discs were applied to the agar surface. A control disc, contained DMSO only was included. Plates were incubated for 48 hours at 42 °C under microaerobic conditions and the diameter of the growth inhibition zones measured.

2.17. Tricine-SDS-PAGE for \textit{C. jejuni} capsule and LOS visualisation.
\textit{C. jejuni} was grown on CBA plates as previously described (Section 2.4). Cells were harvested and suspended in PBS to an OD\(_{600}\) of 1. The material was centrifuged at 15,000 g and the supernatant removed. Pellets were resuspended in 300 µl 1x Tricine-SDS sample buffer (100 mM Tris pH 6.8, 5 % (w/v) SDS, 10 % (v/v) glycerol and 0.025 % (w/v) Bromophenol blue) by heating to 95 °C for ten minutes. Proteinase K (Bio-Rad), at 1 mg/ml concentration, was added and the sample incubated at 55 °C for two hours. The supernatants were loaded onto a precast 12.5 % Tricine-SDS-PAGE gel
along with broad range pre-stained markers. Electrophoresis was performed as previously described (Section 2.8), in an XCell Surelock™ Mini-Cell with 1x anode (200 mM Tris-HCl at pH 8.9) and 1x cathode (100 mM Tris, Tricine, 0.1 % SDS) running buffers, and was carried out at 4 °C [180]. Initially electrophoresis was done at 40 V, then the voltage was then increased to 125 V until the dye front ran approximately two thirds of the way through the gel.

For capsule staining, gels were washed three times for seven minutes in fixing buffer (55:40:5, H₂O:ethanol:acetic acid) and then incubated in a 0.5 % (w/v) solution of Alcian Blue 8X (Sigma Aldrich, USA) stain for two hours. Gels were washed with fixing buffer to remove excess stain.

LOS staining was based on previous work [181]. Following electroporation gels were washed twice for 30 minutes in fixing buffer as described above. The gel and its contents were then oxidised for five minutes in a solution of 0.7 % (w/v) periodic acid in fixing buffer, and were washed three times in dH₂O for 15 minutes. Silver stain was prepared as follows. Exactly 1 ml of ammonium hydroxide (14.5 M) was added to 14 ml of 0.1 M sodium hydroxide and to this 2.5 ml of a 20 % silver nitrate solution was slowly added with constant agitation. Using dH₂O the total volume was increased to 75 ml. Gels were immersed in the silver stain solution for ten minutes, washed three times with dH₂O for 15 minutes, and then immersed in developer (50 mg l⁻¹ citric acid and 0.5 ml l⁻¹ of 37 % formaldehyde). Staining was halted with the addition of a 5 % CH₃CO₂H solution.

2.18.1. Fourier Transform Infrared (FT-IR) Spectroscopy

*Campylobacter* strains and species were grown on CBA plates, containing no antibiotics, for 24 hours under microaerobic conditions at 42 °C. From each plate, a single colony of cells was used to inoculate a fresh CBA plate and bacteria were incubated as described above. Following a further round of single colony isolation/growth, cells were harvested from the plate surface using a plastic loop and suspended into 500 µl of sterile saline solution (0.9 % NaCl). The suspension was
centrifuged, washed three times in this solution and 20 µl of each sample was added, in triplicate, onto a 96 well silicon plate. A blank disc was used to produce a baseline spectrum. The plate was then dried in an oven at 50 °C for 20 minutes.

The plate was loaded onto a motorised microplate module HTS-XT attached to an Equinox 55 FT-IR spectrometer (Bruker Spectrospin Ltd., Coventry UK). The motorised module of this instrument introduced the plate into the airtight optics of the instrument, in which tubes of desiccant are contained to remove moisture [182]. A deuterated triglycine sulfate (DTGS) detector was employed for transmission measurements of the samples to be acquired. Spectra were collected every minute using a scan range of 4000–600 cm\(^{-1}\) and a resolution of 4 cm\(^{-1}\). The data were collected using OPUS v 4.0 software supplied by the manufacturers (Bruker Optics), and then exported as ASCII files, which were subsequently imported into Matlab version 6.1. (The Math Works, Inc., Natick, MA, USA) for data analysis \(\text{vide infra}\). The maintenance and running of the FT-IR machine was carried out by Dr David Ellis (University of Manchester).

2.18.2. Cluster Analysis of FT-IR
The dimensions of the data set were reduced by the use of principal component analysis (PCA) [183]. Matlab was used to perform PCA analysis according to the NIPALS algorithm [183]. Plots based on the separation of the first two principal components produced the best 2D representation of the natural variance of the data. Discriminant function analysis (DFA) was then used to separate the cell types based on the principal components and the use of prior knowledge of which cell types produced which spectra. The implementation of DFA followed Fisher's linear discriminant analysis. Given \(n\) samples belonging to \(g\) groups (classes) it looked for the linear function \(f(x)\) maximizing the ratio of the between-groups sum of squares to the within-groups sum of squares. The data were then plotted to determine the resolution of the different cell types. Analysis was performed by Dr Elon Correa (University of Manchester).
2.19. Surface Enhanced Raman Spectroscopy (SERS)

*C. jejuni* cells were grown on CBA plates for 24 hours under microaerobic conditions as described previously. Following two rounds of single colony isolation and subculture, cells were harvested using a sterile loop and suspended into 500 µl of sterile saline solution (0.9 % NaCl). Cell suspensions were centrifuged, washed three times in this solution and 6 µl of the cell suspension was mixed with 4 µl of a silver colloid solution, prepared as previously described [184]. The sample was then applied to a 15 mm diameter polished CaF<sub>2</sub> disk.

SERS spectra were collected on a Renishaw 2000 Raman microscope (Renishaw Plc, Wotton-under-Edge UK) with a low-power (24 mW) near infrared 785 nm diode laser [184]. The machine was run using the GRAMS WiRE software package (Galactic Industries, USA) and the instrument was calibrated using a silicon disc as a static spectrum at 520 cm<sup>-1</sup> for 1 s. Sample discs were focused under a 50x objective lens and 50 random points chosen to collect 50 spectra.

The data were analysed in Matlab and were baseline corrected, filtered and normalised. Robust PCA was applied and any strong outliers removed. PCA was then performed on the remaining samples. The running and maintenance of the SERS equipment was carried out by Mr Samuel Mabbot and the PCA analysis run by Dr Yun Xu (University of Manchester).
2.20. Bacterial cell strains

Seven strains of *Escherichia coli* were used in this project (Table 2.5) and were all acquired from Novagen.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype/description</th>
</tr>
</thead>
</table>
| BL21 (DE3)          | $F^-\text{ompT}\,hsdS\,\begin{smallmatrix}r^+ \\
B_B \end{smallmatrix}\,m^-\begin{smallmatrix}r^+ \\
B_B \end{smallmatrix}\,\text{gal}\,dcm\,(DE3)$                            |
| HMS174 (DE3)        | $F^-\text{ompT}\,hsdR\,\begin{smallmatrix}r^+ \\
k_{12} \end{smallmatrix}\,m^+\begin{smallmatrix}r^+ \\
k_{12} \end{smallmatrix}\,(DE3)\,(Rif^R)$ |
| Rosetta (DE3)       | $F^-\text{ompT}\,hsdS\,\begin{smallmatrix}r^- \\
B_B \end{smallmatrix}\,m^+\begin{smallmatrix}r^- \\
B_B \end{smallmatrix}\,\text{gal}\,dcm\,(DE3)\,\text{lacY1}\,pRARE2\,(Cm^R)$ |
| Rosetta-Gami (DE3)  | $\Delta\text{ara-leu}\,7697\,\Delta\text{lacX74}\,\Delta\text{phoAPvull}\,\text{phoR}\,\text{araD139}\,\text{ahpC}\,\text{galE}\,\text{galK}\,\text{rpsL}\,(DE3)\,F^+\,[\text{lac}^+\,\text{lacI}^q]\,\text{pro}\,\text{gor522::Tn10}\,\text{trxB::kan}\,pRARE\,(\text{Cm}^R,$ |
|                     | $\text{Kan}^R,\,\text{Str}^R,\,\text{Tet}^R)$                                               |
| Origami B (DE3)     | $F^-\text{ompT}\,\text{hsdSB}\,\begin{smallmatrix}r^- \,m^+ \\
B_B \end{smallmatrix}\,\text{gal}\,\text{dcm}\,\text{lacY1}\,\text{ahpC}\,\,(DE3)\,\text{gor522::Tn10}\,\text{trxB}\,(\text{Kan}^R,\,\text{Tet}^R)$ |
| XL1-Blue            | $\text{supE44}\,\text{hsdR17}\,\text{recA1}\,\text{endA1}\,\text{gyrA46}\,\text{thi}\,\text{relA1}\,\text{lac}\,\text{F}^+\,[\text{proAB}^R\,\text{lacI}^q\,\text{lacZM15}\,\text{Tn10}]\,(\text{Tet}^R)$ |
| NovaBlue            | $\text{endA1}\,\text{hsdR17}\,\begin{smallmatrix}(r_{k_{12}} \,m_{k_{12}}) \\
\text{K12} \end{smallmatrix}\,\text{supE44}\,\text{thi-1}\,\text{recA1}\,\text{gyrA96}\,\text{relA1}\,\text{lac}\,\text{F}^+\,[\text{proA}^R\,\text{B}^R\,\text{lacFZM15:\,Tn10}]\,(\text{Tet}^R)$ |
| DH5α                | $F\,\text{endA1}\,\text{glvV44}\,\text{thi-1}\,\text{recA1}\,\text{gyrA96}\,\text{deoR}\,\text{nupG}\,\text{Φ80dlacZ}\,\Delta\text{M15}\,\Delta[\text{lacZYA-argF\,U169}\,\text{hsdR17}(r^{\text{p}}\,\text{m}^{\text{p}}\,\text{K})],\,\lambda^-$ |

Table 2.5. *Escherichia coli* strains used in this study. *E. coli* strains used for molecular biology and gene expression/protein production work done in this project are listed, along with their genotypes.

2.21. Restriction enzyme digests

Using the appropriate restriction enzymes, a gene of interest was cut out from its parent vector. Plasmid DNA (~300 ng) was mixed with 4 µl of a commercial, 10x reaction buffer (NEB, USA). 1.5 µl (30 U) of the required restriction enzyme was added and the total volume made up to 40 µl with sterile dH2O. The solution was incubated at 37 °C for two hours with frequent mixing.

The digested DNA was mixed with a 6x concentrated bromophenol loading dye (0.25 % bromophenol blue, 0.25 % xylene cyanol FF, 30 % glycerol in dH2O) before being electrophoresed on a 0.8 % agarose gel. The gel was cast using the following method.
To 100 ml of a Tris acetate EDTA buffer (TAE) (40 mM Tris, 20 mM acetic acid, 1 mM EDTA at pH 7.0) 800 mg of agarose (Fisher Scientific, UK) was added and dissolved under heating. At a sufficiently cool temperature, 5 µl of ethidium bromide (10 mg/ml stock) was added. The agarose was poured into a casting tray with combs and allowed to set. The gel was placed in an Owl EasyCast B1 Mini Gel Electrophoresis Systems (Thermo Scientific, UK), immersed in the TAE buffer and the samples along with a 2-log DNA ladder (NEB) loaded. Electrophoresis was carried out at 100 V for 45-60 minutes. Following electrophoresis, the gel was visualised using a GENE FLASH Syngene Bio imager and an image captured on a Computar H6Z0812 lens camera connected to a Sony UP895MD printer.

As required, the agarose was then viewed on a transilluminator and the appropriate band(s) cut out using a scalpel. The DNA from the gel slice was then extracted using a Qiagen gel extraction kit under the manufacturer’s instructions. The DNA was eluted in 30 µl of EB buffer (10 mM Tris-HCl, pH 8.5) prior to estimation of its concentration (ng/µl) using a Thermo Scientific ND2000 NanoDrop instrument.

### 2.22. *E. coli* transformations

Competent *E. coli* cells were grown in a 5ml LB O/N starter culture with agitation at 37 °C, using appropriate antibiotics where required. A fresh culture was prepared from this initial culture using a 1 % (v/v) inoculum and grown under the same conditions until an OD₆₀₀ of ~0.5 was achieved. 1 ml aliquots were removed and pelleted using a bench top centrifuge at 17,000 g for two minutes. The supernatant was removed and the pellet resuspended in 500 µl of ice cold, sterile 50 mM CaCl₂. The cells were kept on ice for 30 minutes. 1 µl of plasmid was added (typically 25 ng), carefully mixed with the cells and the mixture kept on ice for another 30 minutes. The cells were heat shocked at 42 °C for 40 seconds and put back into ice for a further two minutes. 800 µl of SOB was then added. With the exception of HMS174 (DE3) cells, which are a fast growing cell line, transformed cells were incubated for one hour at 37 °C with shaking. Transformants were plated out onto LB agar plates, with appropriate antibiotics present, and allowed to grow O/N in a 37 °C oven.
Single colonies of non-expression host strain transformants were used to inoculate 5 ml LB starter cultures (as previously described) with the appropriate antibiotic added. 1 % (v/v) of this starter culture was used to inoculate a fresh 5 ml LB culture, which was grown at 37 °C until an OD$_{600}$ of 0.6 was reached. The culture was then centrifuged at 3000 g for five minutes in a Heraeus Labofuge 400R centrifuge. Plasmid DNA was recovered from the cell pellet using a Qiagen mini-prep spin kit (Qiagen, West Sussex UK) following the manufacturer’s protocol. The recovered plasmid was used as a stock for transformations into expression systems and kept at -20 °C. The presence of the gene of interest was confirmed by restriction digestion as previously described. Sequencing of the entire genes (Eurofins, MWG Operon) was performed to confirm correct DNA sequences.

2.23. Plasmid ligation

A Quick ligation kit (NEB, USA) was used to recircularise digested plasmids. 100 ng/µl DNA was mixed with 2x reaction buffer and 1 µl ligation buffer added to make a final volume of 20 µl. The ligation reaction proceeded at 25 °C for five minutes. The product was then used immediately for transformation into non expression strain *E. coli* cells and plasmid stocks were created as described previously.

2.24. Site Directed Mutagenesis (SDM)

A QuikChange II site directed mutagenesis kit (Stratagene) was used to mutate the *cj1411c* gene using the manufacturer’s protocol. Oligonucleotide primers (Table S1) for introduction of specific mutations were designed using the Agilent QuikChange Primer design software tool and then synthesized by Eurofins MWG Operon (Ebersberg, Germany). The mutation was introduced using a thermal cycling reaction (Table 2.6) using the reagents described in the protocol. Following PCR, the template plasmid was digested using 1µl DpnI (10 U) at 37 °C for two hours. The digested plasmids sample was then transformed into competent NovaBlue *E. coli* cells on antibiotic selective plates, and transformants grown overnight at 37 °C. Plasmids from colonies were
extracted (as previously described) and sent for sequencing to confirm the presence of the desired mutation in the gene, and to check that no other mutations had occurred.

<table>
<thead>
<tr>
<th>Segment</th>
<th>Cycles</th>
<th>Temperature (°C)</th>
<th>Time (mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>95</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>95</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 2.6. PCR cycling parameters for SDM. The table lists the temperatures and times used when generating new mutants of the cj1411c gene by PCR reaction.

2.25. Preparation of E. coli glycerol stocks

E. coli cell line stocks were prepared by inoculating 5 ml LB with a single colony from a plate of a desired transformant. The cells were grown at 37 °C until the culture reached an OD$_{600}$ of 0.5-0.7. 200 µl of the culture was then added to 800 µl of sterile 80% (v/v) glycerol, mixed and stored at -80 °C. Stocks were recovered by streaking thawed samples out onto antibiotic selective LB agar plates.

2.26. E. coli expression trials of the cj1411c gene product

Individual colonies from each cell strain/transformant were picked from plates and inoculated into 5 ml of fresh LB containing Carb (50 µg/ml). In total, four replicates of each strain were prepared. These were grown to an OD$_{600}$ of ~0.6. For half of these cells 1 mM IPTG was used to induce cj1411c gene expression, and all cells were then incubated at either 25 °C or 37 °C. Cell samples were taken at time points, $t = 0, 3, 5$ and 16 hours. Optical densities were taken at each time point to measure growth rates. Cells were centrifuged in a bench top centrifuge at 17,000 g for one minute and the pellet resuspended in TE buffer (50 mM Tris, 1 mM EDTA, pH 7.2) for SDS-PAGE analysis as previously described. Cell samples were prepared in a consistent fashion throughout the study (Table 2.7).
Table 2.7. Cell sample preparation for SDS-PAGE analysis. The general preparation regime for all cell samples used in protein expression trials by SDS-PAGE is given.

### 2.27. CYP172A1 solubility trials

1x Bugbuster (Novagen) reagent, dissolved in TE buffer containing DNase (0.5 μg/ml), was used to resuspend the cell pellets from 1 ml aliquots of *E. coli* cell samples and the mixture was shaken gently at 22 °C for 30 minutes. A small volume of the sample was then removed (to analyse total cell protein). The remaining cell suspension was centrifuged at 17,000 g for five minutes in a bench top centrifuge. The supernatant from this sample (soluble protein) was separated from the pellet (insoluble protein), with the latter suspended in TE buffer (200 µl). SDS-PAGE analysis was conducted by running the protein samples (prepared as in Table 2.7) on 10 or 12 % precast gels.

### 2.28. Temperature trials for CYP172A1 expression in *E. coli*

*E. coli* strains HMS174 (DE3) and Rosetta (DE3) containing the pET15b/cj1411c construct were grown in the presence of Carb (50 μg/ml) at 37 °C until an OD<sub>600</sub> of 0.6 was reached. Prior to gene induction with 1 mM IPTG, the cells were cooled to 25 °C or 18 °C. Controls with no IPTG induction were also used. Time points for cells grown at all three temperatures were taken at \( t = 0, 3, 5 \) and 16 hours. Samples were then analysed by SDS-PAGE as before. A 1x Bugbuster solution was used to check solubility levels of CYP172A1, as previously described.
2.28.1 IPTG gene induction trials
Five flasks of *E. coli* HMS174 (DE3) cells transformed with the pET15b/cj1411c plasmid construct were grown at 37 °C with Carb (50 μg/ml) to an OD<sub>600</sub> of 0.6. All the cultures were cooled to 23 °C and then gene expression was induced with varying concentrations of IPTG, ranging from 0 to 0.5 mM. Cell samples for analysis were removed at time points at \( t = 0, 2 \) and 16 hours. Expression levels of CYP172A1 were then analysed by SDS-PAGE as previously described.

2.28.2. Growth Medium trials for target P450 production
Three growth medium types, Terrific broth (TB), Luria Bertani (LB) and 2x Yeast extract Tryptone (YT) were used as media for growth of *E. coli* HM174 (DE3) transformed with plasmid constructs in the presence of Carb (50 μg/ml). At an optimum OD<sub>600</sub> of ~0.6, gene expression was induced with 1 mM IPTG after cooling the cultures to 23 °C. Time points were taken at \( t = 0 \) and 16 hr. P450 expression at these time points was analysed by SDS-PAGE and cells were then treated using 1x Bugbuster to analyse solubility levels of CYP172A1, as previously described.

2.29. Expression and purification of CYP172A1
CYP172A1 was produced in *E. coli* HMS174 (DE3). Cultures were grown at 37 °C in 2xYT medium with Carb (50 μg/ml) until an OD<sub>600</sub> between 0.6 and 0.8 was reached. Cells were cooled to 23 °C and gene expression was induced with 1 mM IPTG. Cells were then grown for a further 16 hour period at the same temperature prior to collection by centrifugation (5000 g for seven minutes at 4 °C). Cells were washed and resuspended in a minimal volume of 50 mM potassium phosphate buffer (KP<sub>i</sub>), containing 250 mM KCl, 5 mM imidazole and 10 % (v/v) glycerol at pH 8.0. To the cell suspension, complete EDTA free protease inhibitor tablets (Roche) were added (one tablet/200 ml cell suspension) Cell lysis was achieved by 2 rounds of cell breakage using a continuous flow French Press at 950 lb/in<sup>2</sup> with the material kept cold with ice. Lysed cell debris was removed by ultracentrifugation (40,000 g for 35 minutes at 4 °C).
and the supernatants pooled. The solution was mixed with a Ni-NTA (Qiagen, West Sussex, UK) slurry O/N, pre equilibrated in the same buffer with the addition of 10 mM imidazole. The mixture was poured into an empty 70 ml plastic column (Thomson Instrument Company, California, USA) and the unbound material eluted. The column material was subjected to increasing concentrations of imidazole and the protein eluted with 250 mM imidazole. The eluate was concentrated and buffer exchanged in Amicon Spin filters with a molecular weight cut off of 30 kDa. The new buffer consisted of either 100 mM Tris with 250 mM KCl at pH 7.5, or 50 mM KP₁ with 200 mM KCl at pH 7.5. At a volume of less than 3 ml, CYP172A1 was further purified by Sephacryl S200 gel filtration using an AKTA purifier. For large protein volumes a Hi Load Superdex 26/60 column was pre-equilibrated with same buffer used for the protein, and 4 ml fractions were collected. For smaller volumes, a Tricorn Superdex 10/30 column pre-equilibrated with protein buffer was used with 0.5 ml fractions collected. Fractions were analysed by SDS-PAGE and UV-visible spectroscopy to determine their purities and P450 content. Highly pure proteins were then reloaded onto the same column for further purification with identical analysis of fractions. Finally, selected fractions were concentrated in Amicon Spin filters with a cut off of 30 kDa, to a final volume of less than 1 ml and either used immediately or stored in a -80 °C freezer in 30 % (v/v) glycerol.

2.30. CYP172A1 antibody production
An antibody specific to CYP172A1 was produced using a rabbit SPEEDY 28 day immunization protocol (Eurogentec, Belgium). In brief purified fractions of CYP172A1 were prepared in 100 mM Tris with 200 mM KCl at pH 7.5, at a concentration of 0.4 mg/ml. Over the course of 18 days two rabbits were each injected with 400 µg of protein. On days 0, 18 and 21 the rabbits were bled and the serum kept as pre-immune, middle and final bleeds, respectively. These sera where then used as the primary antibody for western blotting.
2.31. UV-Visible spectroscopy

Unless otherwise stated all UV-Visible spectroscopic measurements were obtained using a Cary 50 UV-Visible spectrophotometer (Varian, UK). Either a 1 or 0.01 cm path length quartz cuvette was used for measurements. Optical titrations were typically carried out at 22 °C with CYP172A1 at concentrations between 0.5 and 5 μM. The absorbance was recorded with a wavelength scan between 800-250 nm. Purity (Reinheitzahl, Rz) ratios were obtained by comparing the ratios of $A_{418}$ at the heme Soret peak and $A_{280}$ at the protein peak. Concentrations of CYP172A1 were determined by measuring the absorbance at the Soret (heme) peak for the oxidised (ferric) low-spin form of hemoprotein, and using the calculated molecular extinction coefficient $\varepsilon_{419} = 105 \text{mM}^{-1} \text{cm}^{-1}$.

2.31.1. Dual beam UV-Visible spectroscopy

For compounds which have a strong UV-Visible spectrum in their own right (and thus interfere with the P450 heme spectrum) a dual beam Cary 300 UV-Visible spectrophotometer (Varian, UK) was used. Optical titrations were carried out using 1 cm quartz cuvettes that were blanked against each other across the required wavelength using the same buffer. The reference sample contained the same experimental set up except for the omission of CYP172A1 enzyme. The ligand was then added to both the reference and experimental cuvette in identical amounts prior to measurements being taken. In this way, absorbance changes specific to the interaction between the compound of interest and CYP172A1 could be measured, with non-specific absorbance contributions from the compound cancelled out.

2.31.2. Fluorescence spectroscopy

Tryptophan emission spectra from CYP172A1 exposed to increasing concentrations of guanidine hydrochloride (GdnHCl) were measured using a Cary Eclipse fluorescence spectrometer (Varian, UK). CYP172A1 was prepared in 10 mM Tris buffer with 150 mM KCl at pH 7.5. A 6 M GdnHCl solution was prepared in the same buffer and
increasing concentrations added to the protein sample. Five minute equilibration times were allowed before a fluorescence reading was taken. The fluorescence emission spectrum between 300-500 nm was measured with an excitation wavelength of 280 nm. The excitation/emission slits were set to 5 nm on the instrument. The emission wavelength maximum and intensity were recorded at each [GdnHCl] and the percentage increase in emission plotted against [GdnHCl] in order to determine the extent of unfolding of the protein based on the change in tertiary structure through perturbation of tryptophan fluorescence.

2.32. Pyridine Hemochromogen Assay
A pyridine hemochromogen assay was carried out to quantify the heme content in P450 samples and to enable the determination of the molar extinction coefficient for the oxidized P450 heme at its Soret peak. Data were analysed using the Berry and Trumpower method [185]. A stock solution containing 0.8 mM potassium ferricyanide (K₃Fe(CN)₆), 0.2 M NaOH and 40 % pyridine was prepared and 500 μl of this solution added to 500 μl of CYP172A1 (10 μM). A UV-Visible spectrum of the oxidised protein mixture was taken. The heme was then reduced using a minimum amount of solid sodium dithionite. The concentration of heme present was then 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. The extinction coefficient for CYP172A1 at the heme Soret peak (418 nm) was then calculated using the Beer-Lambert law. Multiple samples were analysed to ensure consistency of the calculated molecular extinction coefficients at the oxidized CYP172A1 Soret peak [186].

2.33. Redox potentiometry
Spectroelectrochemical studies were performed in order to determine the midpoint reduction potential of the CYP172A1 heme iron. Experiments were conducted anaerobically in a glove box (Belle Technology, Weymouth, UK). All solutions were
degassed under N\textsubscript{2} gas prior to use. Concentrated samples of CYP172A1 were passed down an anaerobic Sephadex G25 column (inside the glove box) to remove residual oxygen. The column was equilibrated using anaerobic 100 mM KP\textsubscript{i} buffer with 250 mM KCl at pH 7.5. A 4.99 ml protein solution (5-10 µM) was put into a cooled jacketed reaction vessel set at 15 °C. Mediators (Table 2.8) were added to give a final volume of 5 ml and to facilitate transfer of electrons between the protein and electrode. The vessel was placed on a magnetic stirrer and the solution gently agitated using a magnetic flea. A Cary 50 Bio UV-Visible spectrophotometer (Varian, UK) was used to monitor changes in absorbance via a fibre optic probe immersed in the protein solution. A potentiometer (Mettler Toledo, Columbus, USA) with reference probe was immersed in the solution and used to monitor the potential during reduction and oxidation of the protein.

<table>
<thead>
<tr>
<th>Mediator</th>
<th>Experimental concentration (µM)</th>
<th>Electrode Potential (mV vs. SHE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxynaphthoquinone (HNQ)*</td>
<td>0.3</td>
<td>-145</td>
</tr>
<tr>
<td>Phenazine methosulfate (PMS)*</td>
<td>2</td>
<td>+80</td>
</tr>
<tr>
<td>Benzyl viologen (BV)</td>
<td>1</td>
<td>-311</td>
</tr>
<tr>
<td>Methyl viologen (MV)</td>
<td>0.3</td>
<td>-430</td>
</tr>
</tbody>
</table>

Table 2.8. List of mediators used in redox potentiometry. The respective concentrations of the mediators used as well as their redox potentials (relative to the standard hydrogen electrode SHE) are given. *indicates that these compounds were made fresh for each experiment.

UV-Visible absorption spectra (250-800 nm) of the CYP172A1 sample were recorded before and after mediator addition. The protein was reduced by addition of small volumes (~0.2 µl) of a concentrated sodium dithionite solution according to Dutton’s method [187]. After each addition a five minute equilibration period was allowed before a spectrum was taken along with the respective reading on the potentiometer. This process was repeated until the P450 was fully reduced and further additions of reductant caused no spectral changes. Reoxidation of the P450 was achieved by the addition of a concentrated potassium ferricyanide solution (as described for dithionite in the reductive
direction) and spectra and potentials recorded as before. Data were analysed using Origin software (OriginLab, Northampton MA, USA) and changes in the absorbance were plotted against the applied potential (corrected versus SHE by adding 244 mV to the observed potential). Data were then fitted using the Nernst function (Equation 1), in Origin 7.0, to determine the midpoint potential for the heme iron Fe(III)/Fe(II) couple.

\[(AABS + BABS \times 10^{(E^0 - x)/C})/(1 + 10^{(E^0 - x)/C})]\]

Equation 1: AABS represents the oxidised absorbance coefficient, BABS represents the reduced absorbance coefficient, \(E^0\) represents the midpoint potential, \(x\) is the applied potential, \(C\) is a constant (59 mV for one electron reduction at 298 K) derived from the gas constant (\(R=8.315 \text{ J mol}^{-1} \text{ K}^{-1}\)), absolute temperature (in K) and Faraday’s constant (96500 C mol\(^{-1}\)) to generate the term \(RT/nF\), where \(n\) is the number of electrons.

### 2.34. Differential Scanning Calorimetry (DSC)

A DSC baseline was produced by using a 10 mM Tris buffer with 100 mM KCl at pH 7.5, which was the same buffer in which the protein was prepared. The buffer was degassed prior to use. The baseline was collected from 20 °C to 90 °C at a rate of 60 °C/hour. A protein sample of 0.5 mg/ml was prepared in the same buffer and run under the same conditions. All thermal transitions were record by a VP-DSC instrument (MicroCal Inc., Amherst, MA, USA).

### 2.35. Electron Paramagnetic Resonance (EPR) Spectroscopy

EPR protein samples were prepared at a concentration of 200 µM in 300 µl of 100 mM Tris buffer with 150 mM KCl at pH 7.5. Samples containing P450 ligands were prepared with ligands at 10x their \(K_d\) value (as determined by optical titration) added to the P450 sample. The solutions were transferred into EPR tubes (0.4 mm o.d Suprasil quartz (Wilmad)) using a 1 ml syringe with attached cannula. Samples were frozen in liquid nitrogen and stored at this temperature until spectra were acquired. EPR spectra were recorded at X-band (~9.4 GHz) on a Bruker ER-300D series electromagnet with a
microwave source interfaced with a Bruker EMX control unit (Oxford Instruments, UK). The temperature was maintained at 10 K using a helium flow cryostat coupled with an ITC 503 controller (both Oxford Instruments). EPR spectra were collected using a microwave power of 0.5 mW, the modulation frequency 100 KHz and modulation amplitude of 5 G. All g values were calculated using the software package supplied with the instrument. EPR spectra were collected by Drs Stephen Rigby and Karl Fisher (University of Manchester).

2.36. Circular Dichroism (CD) spectroscopy

CD spectra were recorded in the far UV (190-260 nm) spectral region at 22 °C using a Chirascan CD spectrophotometer (Applied Photophysics, UK) with an incorporated Chirascan Pro-Data software package. Spectra for CYP172A1 (0.75 μM) in 10 mM Tris buffer with 100 mM KCl at pH 8.0 were recorded using a quartz cell of 0.1 mm path length at scan rates of 10 nm/min. Secondary structural content was analysed using the K2d program (http://kal-el.ugr.es/k2d/spectra.html).

2.37. Multiangle Laser Light Scattering analysis

Polydispersity of P450 samples were measured using Multiangle Laser Light Scattering (MALLS) analysis. Samples of purified recombinant protein were applied to a Superdex 200 gel filtration column (GE Healthcare) running at a flow rate of 0.71 ml/min at concentrations ranging from 250-500 μg/ml. Samples eluting from the column passed through an in-line DAWN HELEOS-II laser photometer (laser wavelength 658 nm) and an Optilab rEX refractometer with a QELS dynamic light scattering attachment. Light scattering intensity and eluant refractive index (concentration) were analyzed using ASTRA v5.3.4.13 software to give a weight-averaged molecular mass (Mw). Maintenance and use of the light scattering instrument was performed by Ms. Marj Howard (University of Manchester).
2.38. Thermofluor assay for P450 protein stability analysis
Thermofluor assays were performed in 96 well RT-PCR plates on a Bio-Rad CFX96 real time system C100 thermal cycler (Bio-Rad Laboratories, Hercules, USA). Measurements were performed using an excitation wavelength of 490 nm and emission wavelength of 575 nm. Unfolding curves were generated by use of a gradient between temperatures of 15 °C to 95 °C. A fluorescence measurement was taken after every 0.2 °C increase in sample temperature. The change in fluorescence/temperature was plotted against temperature to estimate a melting temperature ($T_m$). These data were plotted and analysed using Origin software.

A JBS buffer screen with or without additives was conducted to indentify P450 stabilising conditions. A final protein concentration of 6 mg/ml was found to be optimal when mixed with 7.5 µl of SYPRO orange (final concentration 23.5x). The 100 mM buffers from the JBS solubility Kit (Jena Bioscience, Germany) (Table 2.9) were added at a final concentration of 50 mM to the protein/SPYRO mixture to make a final volume of 25 µl.

Buffers which produced the highest $T_m$’s were selected for a secondary screen using the additive screen. A final protein concentration of 6 mg/ml was mixed with 2.5 µl of SYPRO orange (final concentration 23.5x). The 100 mM buffer(s) from the JBS solubility Kit were added at a final concentration of 50 mM to the protein/SPYRO mixture. Finally, 5 µl of the additive was included (Table 2.10) to make a final volume of 25 µl. Further additives screens were investigated with either econazole or 1-phenyl imidazole bound to the protein. The selected ligand was pre-bound to the protein in the protein stock solution. The binding of the ligand was observed by spectrophotometry to ensure complete binding and an appropriate concentration of ligand then added to ensure saturation prior to the screen. The samples were then prepared as previously described.
### Table 2.9. List of buffers present in JBS buffer screen and their respective pH values.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine</td>
<td>3.0</td>
</tr>
<tr>
<td>Citric acid</td>
<td>3.2</td>
</tr>
<tr>
<td>PIPPs (Piperazine-1,4-bis(propanesulfonic acid)</td>
<td>3.7</td>
</tr>
<tr>
<td>Citric acid</td>
<td>4.0</td>
</tr>
<tr>
<td>Sodium acetate</td>
<td>4.5</td>
</tr>
<tr>
<td>Sodium/potassium phosphate</td>
<td>5.0</td>
</tr>
<tr>
<td>Sodium citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>Sodium/potassium phosphate</td>
<td>6.0</td>
</tr>
<tr>
<td>Bis-Tris</td>
<td>6.0</td>
</tr>
<tr>
<td>MES (2-(N-morpholino)ethanesulfonic acid)</td>
<td>6.2</td>
</tr>
<tr>
<td>ADA (N-(2-acetamido)iminodiacetic acid)</td>
<td>6.5</td>
</tr>
<tr>
<td>Bis-Tris propane</td>
<td>7.0</td>
</tr>
<tr>
<td>Ammonium acetate</td>
<td>7.0</td>
</tr>
<tr>
<td>MOPS (3-(N-morpholino)propanesulfonic acid)</td>
<td>7.0</td>
</tr>
<tr>
<td>Sodium/potassium phosphate</td>
<td>7.0</td>
</tr>
<tr>
<td>HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)</td>
<td>7.5</td>
</tr>
<tr>
<td>Tris</td>
<td>7.5</td>
</tr>
<tr>
<td>EPPS 3-[4-(2-Hydroxyethyl)-1-piperazinyl]propanesulfonic acid</td>
<td>8.0</td>
</tr>
<tr>
<td>Imidazole</td>
<td>8.0</td>
</tr>
<tr>
<td>Bicine</td>
<td>8.5</td>
</tr>
<tr>
<td>Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>CHES (N-Cyclohexyl-2-aminoethanesulfonic acid)</td>
<td>9.0</td>
</tr>
<tr>
<td>CHES</td>
<td>9.5</td>
</tr>
<tr>
<td>CHAPS (3-[(3-Cholamidopropyl)dimethylammonio]propanesulfonic acid)</td>
<td>10.0</td>
</tr>
</tbody>
</table>
## Additive Stock concentration

<table>
<thead>
<tr>
<th>Additive</th>
<th>Stock concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium chloride</td>
<td>80 mM</td>
</tr>
<tr>
<td>Sodium chloride</td>
<td>200 mM</td>
</tr>
<tr>
<td>Sodium chloride</td>
<td>400 mM</td>
</tr>
<tr>
<td>Glycerol</td>
<td>20 %</td>
</tr>
<tr>
<td>Glycerol</td>
<td>40 %</td>
</tr>
<tr>
<td>CHAPS</td>
<td>8 mM</td>
</tr>
<tr>
<td>Octyl glucopyranoside</td>
<td>0.4 %</td>
</tr>
<tr>
<td>Octyl glucopyranoside</td>
<td>4 %</td>
</tr>
<tr>
<td>Dodecyl maltoside</td>
<td>0.4 %</td>
</tr>
<tr>
<td>Dodecyl maltoside</td>
<td>4 %</td>
</tr>
<tr>
<td>BME (beta mecaptoethanol)</td>
<td>40 mM</td>
</tr>
<tr>
<td>DTT (dithiothreitol)</td>
<td>4 mM</td>
</tr>
<tr>
<td>DTT</td>
<td>20 mM</td>
</tr>
<tr>
<td>TCEP (tris(2-carboxyethyl)phosphine)</td>
<td>120 mM</td>
</tr>
</tbody>
</table>

Table 2.10. List of additives present in JBS buffer screen and their concentrations.

### 2.39. CYP172A1 crystallisation trials

Crystallisation trials were conducted using a small volume 100 nl drops, delivered by a Mosquito pipetting robot (Molecular Dimensions, Newmarket, UK). Protein concentrations ranging from 10 – 25 mg/ml were used and, as required, ligands were added at values of 10x their respective $K_d$’s plus the relevant protein concentration, in order to ensure complete binding. Protein samples were tested using, JCSG-plus, PACT premier Morpheus™ and Clear Strategy (Molecular Dimensions) 96-well crystallography trays. Trays were kept at 4 °C and checked for crystal formation periodically.
2.40. Protein structure modelling
The amino acid sequence of CYP172A1 was aligned with those of other P450’s from the SCOP database. The coordinates from crystal structure of 1JPZ (a N-palmitoylglycine substrate-bound form of the heme [P450] domain of the Bacillus megaterium P450-CPR fusion enzyme P450 BM3 [CYP102A1]) were downloaded into PyMOL and used to construct a homology model of CYP172A1.

2.41. Ligand binding assays
Ligand binding to CYP172A1 was investigated by performing titrations using UV-Visible spectroscopy. Solutions of CYP172A1 (5-10 µM) in a total of 1 ml (in 100 mM tris buffer, with 250 KCl at pH 7.5) were prepared in a quartz cuvette. Stock solutions of concentrated ligands were titrated against the P450 using a Hamilton syringe (1 or 10 µl capacity) to a final volume of ≤ 5 % of the protein solution. After each addition of ligand a spectrum was taken to observe any changes to the spectral features (a sign of ligand binding). The multiple final spectra were analysed using Origin 7 software and difference spectra were generated by subtracting the spectrum for the ligand-free spectrum from each subsequent ligand-bound spectrum in the titration set. The absorbance value at the trough of the difference spectrum was subtracted from the value at the peak (keeping the wavelength pair constant for all spectra analysed in each titration set) to define the overall absorbance difference at each ligand concentration. The absorbance differences were then plotted against the relevant ligand concentration. $K_d$ values were then obtained by fitting $\Delta \Delta$absorbance versus [ligand] using equation 2:

$$A_{obs} = (A_{max} / 2 E_t)(S + E_t + K_d) - \left\{ [(S + E_t + K_d)^2 - (4SE_t)]^{0.5} \right\}$$

Equation 2: $A_{obs}$ represents the change in absorbance at ligand concentrations $S$, $A_{max}$ is the maximum absorbance change at ligand saturation, and $E_t$ is the enzyme concentration.
2.42. Compound Screening

Screening for compounds binding to CYP172A1 was performed by the Screening Unit at the Department of Medicinal Chemistry, Leibniz Institute for Molecular Pharmacology (FMP), Berlin. In brief, screening was done using a medium-throughput pipetting robot (Sciclon 3000, Caliper Life Sciences, Hopkinton, USA), a plate reader (Safire, Tecan, Reading, UK) and robots for washing and dispensing. The FMP library contains 20,000 organic molecules at 10 mM dissolved in DMSO. To 0.4 µl of stock molecules, 40 µl of 10 mM Tris-HCl buffer with 150 mM KCl at pH 8.0 was added to make 100 µM compounds. Compound-specific changes in CYP172A1 absorption spectra (monitored between 310 – 450 nm) were automatically recorded after the addition of 10 µl of CYP172A1 (4 µM). 4-PIM (30 µM) was used as a type II (heme coordinating inhibitor) binding mode reference in these studies.

2.43. P450 CO trapping experiments

A redox system for CYP172A1 was constructed using a three protein system in a 1:2:5 ratio of P450:FldR:FldA/Fdx. Where FldR is *E. coli* flavodoxin reductase, FldA is *E. coli* flavodoxin and Fdx is spinach (sp) ferredoxin. Experiments were carried out in an anaerobic glove box. The reaction buffer (100 mM KP$_1$ buffer, with 250 mM KCl at pH 7.5) was degassed under nitrogen prior to use. Small volumes of concentrated stocks of all the proteins were exposed to the anaerobic environment in the glove box to remove any residual O$_2$. NADPH stock (20 mM) was prepared using the reaction buffer once inside the glove box. The reaction mixture contained CYP172A1 (2 µM), *E. coli* FldR (4 µM), *E. coli* FldA or SpFdx (10 µM), CO gas (saturating), and the reaction was initiated by adding NADPH (100 µM). The compound 204073 (a hit from compound screening) (10 µM) was also added to investigate its effect on electron transfer through the redox partner system. Control experiments were also performed where either the reductase or the FldA/Fdx partner were omitted from the reaction mixture. Spectra (800-250 nm) were recorded every 30 seconds and the formation of a peak at 450 nm was monitored, reflecting formation of the Fe(II)CO complex if the P450, involving rapid binding of CO to the heme iron once it was reduced to the ferrous form.
The reaction was allowed to continue for several hours to allow the reaction to reach near completion. The rate constants for the formation of Fe(II)CO complex under different conditions were determined by plotting the ΔA_{450} vs. time and using a single exponential function (Equation 3) to fit the data in Origin 7.0.

\[
(A \ast (1 - \exp(-(k \ast x))) + b)
\]

Equation 3: A represents the absorbance at the relative wavelength (448 nm), k represents the rate constant for formation, x represents the time (in minutes), and b is an offset value, if the function does not start at zero.

### 2.44. Enzyme turnover studies

Turnover experiments were conducted by incubating CYP172A1 (1 µM) with 200 µM of either compound 204073 or 213071 (substrate-like molecules). 500 µM NAPDH and a NADPH regeneration system, composing of G-6-DH (glucose-6-phosphate dehydrogenase) and G-6-P (glucose-6-phosphate) were added and the reaction volume made up to 5 ml. The reaction was initiated with the addition of NADPH (500 µM) and continued for either two hours at 37 °C or for 16 hours at room temperature with stirring. The sample was recovered by extraction using one volume of dichloromethane (DCM). The reaction mixture was vortexed and then centrifuged with the organic layer being removed. The DCM was removed under vacuum in a Speedy-Vac centrifuge (GeneVac, New York, USA). Dried material was suspended in 200 µl methanol:acetonitrile (50:50, v:v) with 0.01 % formic acid. 20 µl samples were analysed by mass spectrometry using a Waters LCT Mass Spectrometer and a Waters Alliance autosampler, running Waters MassLynx 4.0 software, and run by Dr Reynard Spiess (MIB, University of Manchester). The cone voltage was set to 45 V and source temperature to 80 °C. Controls where CYP172A1 was boiled prior to addition to the reaction mixture were also examined.
2.45. Ion trap liquid chromatography (LC)

Analysis of turnover (TO) products (Section 2.44) was performed on a LC MSD (Agilent 1100 series). The set up contained an autosampler, buffer degasser and DAD (diode array detector) with the LC directly linked to an ion trap mass spectrometer. 10 μl of a 213071 TO reaction was separated on a C18 column (Phenomenex Kinetex, 100 x 4.6 mm) with an inline Krudcatcher (Phenomenex, 0.004 ID) attached. A flow rate of 1 ml/min was used and the column temperature set to 40 °C. The elution profile (Table 2.11) was optimized to give the best separation of individual peaks. The elution of compounds from the column was measured at 280 nm prior to their direct injection into the mass spectrometer. For each identifiable peak a corresponding mass spectrum was produced. Analysis of these compounds and their mass spectra were used to identify the components of the peak.

<table>
<thead>
<tr>
<th>Time (minutes)</th>
<th>% ACN</th>
<th>% Water, 0.1 % TFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>35</td>
<td>65</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>65</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>11</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>13</td>
<td>43</td>
<td>67</td>
</tr>
<tr>
<td>18</td>
<td>43</td>
<td>67</td>
</tr>
<tr>
<td>22</td>
<td>75</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 2.11. Elution profile used to separate 213071 turnover reaction products. The percentage of ACN was increased over a total time of 22 minutes from 35 % to 75 %, with the aqueous phase decreasing from 65 % to 25 %. ACN (acetonitrile), TFA (trifluroacetic acid).
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

3.1. Introduction

*C. jejuni* is a human pathogen that is a leading cause of gastroenteritis world wide, and can even be fatal. Understanding the development of the bacterium as well as the mechanisms of infection is important if successful treatments/vaccines are to be created to limit the levels of infection. The *cj1411c* gene product, CYP172A1, is the sole *Campylobacter jejuni* P450-like enzyme and is largely uncharacterised. The gene is located within the genome of some *Campylobacter* species, and orthologues are also present in the *Helicobacter* species; *H. canadensis, H. pullorum* and *H. winghamensis*, and in the more distantly related *Comamonas testosterone* (Section 1.5). There have been no investigations into the role of this P450 in these species, but the location of *cj1411c*, adjacent to the capsular polysaccharide (CPS) biosynthetic locus, suggests that it may have a role in the generation of surface located features. The CPS is a highly important structural feature of *C. jejuni*, which is involved in the protection from desiccation, avoidance of host immune systems and also in adhesion and virulence. *Campylobacter* and *Helicobacter* species (including *H. pylori*) that lack the P450 gene also lack the full *kps* locus, which is suggestive that the P450 has a biosynthetic role in the generation of the CPS. Although assigning functional roles to genes based on their genomic location can be informative, it can also be misleading and more direct evidence is required [14]. It is therefore important to study the effect of the loss of function of the P450 to investigate if the protein is part of an important biosynthetic role in capsular formation.

Genetic manipulation of *C. jejuni* NCTC 11168H [188] and 81-176 [189] (two well characterised strains) allowed for the creation of P450 mutant and genetically complemented strains (PE1-PE6). These cells were then used to examine the resulting phenotype with the loss of the P450 function in *C. jejuni*. A broad range of experiments were conducted on the PE1-PE6 strains, including examining growth rate, cell
morphology, cell-to-cell interactions and antimicrobial sensitivities. With the emergence of antibiotic resistant strains of *Campylobacter*, it was important to explore the potential of the P450 to provide an antimicrobial resistance function. The azole antifungals, well known for their ability to inhibit P450s *in vitro* were also tested as potential novel antimicrobial agents against *C. jejuni*. If the CPS structure is affected by the loss of P450 function then it would expected that the surface acting antimicrobials would be more effective in the mutant strains and growth defects would be expected if CYP172A1 was essential for cell survival. Due to the genomic proximity of *cj1411c* to CPS biosynthetic genes, gel-based and spectroscopic techniques were used in order to investigate the role of CYP172A1 in the biosynthesis of the CPS. An alteration to the CPS would be detectable, by different staining and spectroscopic profile differences, using both methods and confirm the role of the P450 in *C. jejuni*.

Thus, the overall aim of studies in this chapter was to establish the influence of CYP172A1 on the viability and growth properties of *C. jejuni*, including its susceptibility to P450 inhibitor drugs. Through other microbiological and spectroscopic analysis I sought to define phenomenological and structural changes that might indicate a biochemical function for CYP172A1.

### 3.2. Construction of *cj1411c* insertion knockouts (PE2 and PE5) in *C. jejuni* 11168H and 81-176 strains

Previous work within the laboratory (Jervis and Linton, unpublished data) resulted in the construction of a pGEMTEasy/*cj1411c::kan* plasmid (Appendix, Figure S1.). In this construct, both orientations of the kanamycin resistance cassette (with respect to the *cj1411c* gene) were possible, as the resistance cassette was inserted at a single BclI site within the gene. The presence of the cassette was verified by PCR using primers 375 and 376 (Appendix, Table S1), that annealed to the *cj1411c* gene to produce an amplicon with a predicted size of 2.7 kbp (Figure 3.1). The combination of the 375/513 and the 376/561 primer pairs enabled the identification of a plasmid containing the resistance cassette in the same transcriptional orientation as the *cj1411c* gene. For the primer combination 375/513 and 376/561 predicted product sizes were 0.75 kbp and
0.85 kbp, respectively (Figure 3.1). The plasmid construct used for mutagenesis contained the kanamycin resistance cassette in the same transcriptional orientation as the *cj1411c* gene in order to minimise potential polar effects.

The pGEM/*cj1411c::kan* construct was electroporated into wild type *C. jejuni* 11168H (PE1) and 81-176 (PE4) cells, and potential transformants isolated as described (Section 2.7). Successful homologous recombination and replacement of the native *cj1411c* gene on the chromosome with the version containing the kanamycin resistance cassette was confirmed by colony PCR. Primers 680 and 681 (Appendix, Table S1) were designed to anneal to nucleotides outside of the inserted construct. This was done to confirm that homologous recombination with the chromosome had occurred at the desired chromosomal location. The orientation of the kanamycin resistance cassette, with respect to the *cj1411c* gene, was confirmed using primer combinations 680/513 and 681/561 (Figure 3.2).

PCR amplification of PE1 and PE4 strains, using the same primer combinations as mentioned above, was performed as a control. Following PCR, products were separated on agarose gels to confirm amplicon sizes (Figure 3.3). Using the 680/681 primers wild type (PE1 and PE4) and *cj1411c::kan* insertion knockouts (termed PE2 and PE5 for *C. jejuni* 11168H and 81-176, respectively), produced a product, but this differed in size (1.8 kbp for PE1 and PE4 and 3.3 kbp for PE2 and PE5): as predicted PCR products were not generated for PE1 and PE4 strains using primer combinations 680/513 and 681/561. The orientation of the kanamycin resistance cassette within the *cj1411c* gene in PE2 and PE5 strains was confirmed as primer pairs 680/513 and 681/561 amplified DNA with sizes of 0.9 kbp and 1.1 kbp, respectively. PCR amplification with primer pairs 680/561 and 681/513 failed generate a PCR product as would be predicted.
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

Figure 3.1. PCR verification of the orientation of kanamycin resistance cassette with respect to the *cj1411c* gene in pGEM/*cj1411c::kan* constructs. A) Schematic diagram of *cj1411c* gene (red) with inserted kanamycin resistance cassette (blue). Orientation of the resistance cassette, with respect to the *cj1411c* gene, was checked using specific primer binding sites (as indicated). The expected sizes following PCR amplification (with correct orientation) are displayed. B) Following PCR amplification, products were analysed on a 1 % agarose gel. Band sizes are indicated by a marker (M). For plasmids containing the *cj1411c* gene and resistance cassette in the same orientation, primers 375/376 produced a product of approximately 2.7 kbp (lane 1) Primers 375/513 produced a product approximately 0.75 kbp (lane 2), and primers 376/561 produced a product of approximately 0.85 kbp (lane 3). For plasmids containing the incorrect orientation of the kanamycin cassette, primers 375/376 produced a product of approximately 2.7 kbp (lane 4), but no product was observed from either 375/513 (lane 5) or 376/561(lane 6) primer combinations.
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

Figure 3.2. Schematic diagram of the arrangement of the *cj1411c* gene with the inserted kanamycin resistance cassette. The *cj1411c* gene (red) with inserted kanamycin resistance cassette (blue). Primer binding sites are displayed, along with the expected sizes of DNA products following PCR amplification.

Figure 3.3. PCR verification of homologous recombination event of the pGEM/*cj1411c::kan* plasmid in PE2 (A) and PE5 (B). PCR amplified products were run on a 1 % agarose gel. The band sizes are indicated by a DNA marker (M). DNA amplification using primers 680/681 produced products of approximately 1.8 kbp and 3.3 kbp for the PE1 and PE4 and PE2 and PE5 strains, respectively (lanes 1, 4). Amplified products using primers 680/513 produced a band of approximately 0.9 kbp (lane 5) and primers 376/561 produced a band of approximately 1.1 kbp (lane 6) in PE2 and PE5 only (no product for PE1 and PE4, lanes 2 and 3). Primers 680/561 (lane 7) and 681/513 (lane 8) did not produce an amplified product in PE2 and PE5.
3.3. Construction of the genetically complemented strains (PE3 and PE6) in C. jejuni 11168H and 81-176

Genetic complementation was performed in order to reintroduce a stable copy of the *cj1411c* gene into the chromosome of the insertion knockout mutant strains. To create the complement strains (PE3 and PE6 in strains 11168H and 81-176, respectively) a plasmid construct was created using the pCJ0223 vector, a pUC18 based plasmid containing the *cj0223* pseudogene [190]. This gene was chosen as it is expected to be redundant in *C. jejuni*, and so the insertion of the *cj1411c* gene here would have little or no effect on the cell. The pCJ0223 construct contained a chloramphenicol resistance (cat) cassette [191], which was used as a selective marker for successful homologous recombination. The cat cassette allows for constitutive expression of the downstream gene, in this case *cj1411c*, which should enable to restoration of the gene product to wild type levels. [191]. The *cj1411c* gene and a 30 base upstream region (1,342,472 to 1,343,967 bp of the 11168H genome sequence) were PCR amplified from 11168H genomic DNA using primer pair 1411/1412 (Appendix, Table S1). Primer 1411 contained a NcoI restriction site, and primer 1412 contained a NheI restriction site, which allowed for directional cloning into the pCJ0223 vector. The size of the amplicon was verified by electrophoresis on agarose gel and the amplicon digested with NcoI and NheI. The pUC18/pCJ0223 plasmid containing a cat cassette inserted into the *cj0223* gene was digested with the same restriction enzymes (Figure 3.4). The *cj1411c* gene was inserted downstream of the cat cassette. The resulting plasmid was transformed into *E. coli* DH5α and successful transformants selected on chloramphenicol-containing LB agar plates.

The PCJ0223::*cat/cj1411c* plasmid was sequenced to check that the construct was correct and contained no mutations, and electroporated into PE2 and PE4 cells. Following selection on chloramphenicol-containing plates, individual colonies were subcultured and analysed by colony PCR using the 680/681 and 66/311 primer combinations to check for homologous recombination within the pseudogene of the full length *cj1411c* gene into the host chromosome (Figure 3.5).
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

Figure 3.4. Schematic diagram illustrating the cloning strategy used to generate the *pCJ0223::cat/cj1411c* plasmid. A) Plasmid map of the pCJ0223 with the *cj0223* gene (orange) and inserted cat resistance cassette (green). B) The *cj1411c* gene with an upstream region (red) was ligated into the pCJ0223 plasmid using NcoI and NheI restriction enzymes to create the *pCJ0223::cat/cj1411c* plasmid. Sizes of DNA fragments and plasmids are not to scale. cat, chloramphenicol resistance cassette.

Figure 3.5. Agarose gel electrophoresis of colony PCR products from *cj1411c* C. jejuni strains PE3 and PE6. PCR products were run on a 1% agarose gel and band sizes are indicated by a DNA marker (M). A single colony was subjected to PCR with primers 66/311, to check for the presence of the complement *cj1411c* (1), and 680/681 primers to check for the insertional mutation of the native *cj1411c* gene (2). Band sizes of ~2.5 kbp and 3.3 kbp correspond to the cat regulated *cj1411c* gene and the *cj1411c::kan* mutation insert, respectively.
3.4. Detection of CYP172A1 in *C. jejuni* by probing western blots with CYP172A1 specific antiserum

In order to verify that *cj1411c* is expressed in *C. jejuni* an antiserum against the recombinant expressed protein was raised in a rabbit host (Section 2.30). This antiserum was used as a primary antibody in a western blot to screen for the production of CYP172A1. Cell samples were separated by SDS-PAGE, blotted onto PVDF membranes and probed with the CYP172A1 antiserum and secondary antibody (Section 2.8/9). Initial blots were difficult to interpret due to the generation of multiple bands on the membrane.

In order to improve the immunoblot a cell lysate from the P450 insertional knockout mutant (PE2) was mixed 1:10 with the rabbit antiserum and incubated for 4 hours. This was done in an attempt to remove antibodies to proteins other than CYP172A1 from the cell lysate. This technique has been used successfully to remove cross reactive antibodies generated by rabbit hosts [192]. This preabsorbed antiserum was then used to probe western blots with material from *C. jejuni* strains grown at both 37 °C and 42 °C. This resulted in fewer non specific bands (Figure 3.6). The recombinant His-tagged CYP172A1 was used as a positive control to aid with the detection of CYP172A1 in the *C. jejuni* hosts. The *cj1411c* gene product was detected as a band with a $M_w$ of ~52 kDa in *C. jejuni* strains PE1, PE3, PE4 and PE6 at 42 °C and 37 °C but not in PE2 or PE5 strains. The positive control had a $M_w$ of ~54 kDa (due to the presence of the His$_6$ tag).

These results demonstrated that the *cj1411c* gene product, CYP172A1, is produced during *in vivo* growth in wild type *C. jejuni* (PE1 and PE4). The insertion of a kanamycin cassette results in the loss of the production of the P450 in strains PE2 and PE5. Finally, it was demonstrated that the reintroduction of *cj1411c* into the pseudogene *cj0223* resulted in the production of CYP172A1 to levels similar to those observed in the wild type strains. It was also demonstrated that CYP172A1 is produced at both 37 °C and 42 °C, which indicates that this protein’s expression is not regulated by a thermo switch as suggested for certain other *Campylobacter* proteins [193, 194].
Figure 3.6. Reactivity of *C. jejuni* proteins with preabsorbed CYP172A1 antiserum.

Standardized whole cell lysates of PE1-PE6 cells were separated by SDS-PAGE, transferred onto a PVDF membrane and western blotted. The recombinant protein (100 nM) was used a positive control and was detected with a $M_w$ of ~54 kDa. Cells were grown at both 42°C (A) and 37°C (B) to investigate temperature dependent production of CYP172A1. A band was detected at the expected size of CYP172A1 (~52 kDa) in PE1, PE3, PE4 and PE6, but not in PE2 and PE5. The position of the 50 kDa marker is marked for reference.
3.5. Growth comparisons of *C. jejuni* strains PE1-PE6 under microaerobic conditions at 42 °C

Growth curves of *C. jejuni* strains were generated as described (Section 2.10) to determine if mutagenesis of *cj1411c* affected *in vivo* growth. Cultures of *C. jejuni* were standardised to and OD$_{600}$ of 0.05 in MH, antibiotics were not included to ensure no observed differences were due to different conditions experienced by the cells. Growth was monitored by periodically measuring the OD$_{600}$ of the growing cultures. At three time points (4, 8 and 16 hours) samples were removed and cell viabilities were determined using the Miles and Misra technique [174].

Based on three independent experiments (performed in triplicate), the bacterial cells produced typical growth curve characteristics. There was a lag phase (up to two hours) followed by exponential growth (between 2 to 16 hours) and a stationary phase (Figure 3.7). The OD$_{600}$ values for the PE2 and PE5 strains (red circles) were lower than those observed for the PE1 and PE4 strains (black squares). However, there was no significant difference in viable cell counts (Table 3.1). Discrepancies in the correlation between OD$_{600}$ and viable cells counts have been previously attributed to differences in morphology (spiral vs. coccoid shaped cells) as well as to the release of cell debris [195, 196]. The CYP172A1 complement strains (PE3 and PE6) produced a growth curve with OD$_{600}$ values intermediate between those observed for the wild type and *cj1411c* gene mutants (green triangles). This suggested partial restoration of this phenotypic difference.
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

Figure 3.7. The effect of *cj1411c* insertional knockout and complementation to 24 hour time course growth assays of *C. jejuni* PE1-PE6 strains. A) Comparisons of the growth curves of *C. jejuni* 11168H strains PE1 (black), PE2 (red) and PE3 (green). B) Comparisons of the growth curves of *C. jejuni* 81-176 strains PE4 (black), PE5 (red) and PE6 (green) were monitored by recording the OD$_{600}$ over 24 hours. Error bars represent standard error (SE), n=3. Arrows indicate time points when cell viabilities were recorded.
Viable counts at time points 4, 8 and 16 hours were determined for both the wild type and insertional knockout mutant strains (Table 3.1). The viable counts for all the cells increased from $10^7$ to $10^9$, with doubling times of 1.5 hours for PE1 and PE2 strains and ~1.8 hours for PE4 and PE5 strains. Doubling times of ~2 hours for the 11168 strain have been reported by other groups [197], but for the 81-176 strain longer doubling times of 2.5 hours have been reported [198].

<table>
<thead>
<tr>
<th>Time (hr)</th>
<th>PE1</th>
<th>PE2</th>
<th>PE4</th>
<th>PE5</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$3.8 \times 10^7$ (±0.2)</td>
<td>$2.2 \times 10^7$ (±0.2)</td>
<td>$3.4 \times 10^7$ (±0.9)</td>
<td>$2.0 \times 10^7$ (±0.2)</td>
</tr>
<tr>
<td>8</td>
<td>$2.3 \times 10^8$ (±0.4)</td>
<td>$2.3 \times 10^8$ (±0.2)</td>
<td>$2.2 \times 10^8$ (±0.1)</td>
<td>$1.9 \times 10^8$ (±0.3)</td>
</tr>
<tr>
<td>16</td>
<td>$1.7 \times 10^9$ (±0.2)</td>
<td>$1.9 \times 10^9$ (±0.1)</td>
<td>$1.5 \times 10^9$ (±0.2)</td>
<td>$1.4 \times 10^9$ (±0.1)</td>
</tr>
</tbody>
</table>

Table 3.1. Viable counts of *C. jejuni* 11168H and 81-176 cells over 16 hours of growth. Viable counts of cell strains indicated were measured by the Miles and Misra technique during early exponential phase (4 hours), mid exponential phase (8 hours) and early stationary phase (16 hours). Variation of observed cfu ml$^{-1}$ in brackets, n=3.

### 3.6. Investigating the effect of CYP172A1 on autoagglutination (AAG)

AAG is a rapid method that can be used to examine the ability of *C. jejuni* to interact with surrounding *C. jejuni* cells. Flagella are known to be key mediators in AAG, but other changes to cell surface can also have an affect [54, 175, 199].

AAG data indicated that the *cj1411c* insertional knockout mutant in 11168H resulted in a reduced capacity of the cells to agglutinate (Figure 3.8). From 45 minutes to 120 minutes the observed difference between PE1 and PE2 was statistically significant ($p \leq 0.05$). Restoration of the *cj1411c* gene by complementation partially rescued this phenotype, and from 45 minutes to 120 minutes the observed difference in autoagglutination between PE2 and PE3 was statistically significant ($p \leq 0.05$). A mutant almost completely devoid of flagella, lacking flagellin A protein (*ΔFlaA* [200]), was used as a control, to measure the effect on AAG when fully functional flagella are not present. The result clearly illustrated the significance of intact flagella, as after 120 minutes, where all of the other strains have an OD$_{600}$ below 0.4, the flagella mutant is largely unchanged (OD$_{600} = 0.98$).
In 81-176 strains (PE4-PE6) the loss of the CYP172A1 also resulted in the reduction of the AAG levels when compared to the wild type and complement strains (Figure 3.9). After one hour there were statistically significant differences ($p \leq 0.05$) between the PE4 and PE5 strains. Reintroduction of the *cj1411c* gene resulted in the restoration of wild type like AAG, and from two hours to 18 hours the observed difference in autoagglutination between PE5 and PE6 was statistically significant ($p \leq 0.05$).

The two strains, 11168H and 81-176, produced significantly different aggregation dynamics in the AAG experiments. The 11168H experiment was effectively completed within two hours. Whilst to achieve the comparable levels of AAG in the 81-176 strains, 18 hours was required.

The results obtained for the 81-176 is in agreement with the data analysis of 11168H and further suggests that the mutation causes a change in the charge of the cell surface, which compromised the bacterium’s ability to self-aggregate.
Chapter 3. Investigating the role of the \textit{cj1411c} product, CYP172A1, in \textit{Campylobacter jejuni}.

Figure 3.8. Effect of \textit{cj1411c} insertional knockout and subsequent complementation on AAG activity in \textit{C. jejuni}. A) AAG of PE1, PE2 and PE3 were compared. Cell suspensions were standardized to an OD\textsubscript{600} of 1. PE2 displayed a reduced level of AAG autoagglutination compared to PE1 and PE3. A flagellin A (\textit{ΔFlaA}) mutant was included as a control. Scale bars represent SE, \(n=3\). **denotes a statistical significance between PE1/PE2 and PE2/PE3 using a paired sample students t test (p \leq 0.05).

B) AAG of PE4, PE5 and PE6 were compared. Cells all started with an OD\textsubscript{600} = 1. PE5 displayed a reduced autoagglutination rate compared to PE4 and PE5. Scale bars represent SE, \(n=3\). *denotes a statistical significance between PE4/PE5, ** denotes a statistical significance between PE4/PE5 and PE5/PE6 using a paired sample students t test (p \leq 0.05).
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

### 3.7. The role of CYP172A1 in the morphology of *C. jejuni* examined using electron microscopy

Electron microscopy (EM) is a useful tool when examining the morphology of bacterial cells. Agar grown cells were harvested and stained with uranyl acetate (Section 2.12), which is a negative stain that binds to protein and lipid. One major benefit of uranyl acetate over other stains is that it produces high electron density which improves the contrast, allowing for greater resolution of the image, which is especially useful for small specimens. The stained cells were carefully transferred onto copper grids to prevent shearing of flagella and were viewed using a calibrated TEM machine.

*C. jejuni* 11168H cells displayed typical spiral morphology, with bipolar flagella (Figure 3.10) [77, 201]. The PE1 cells were on average $1.96 \pm 0.76 \mu m$ long and $0.54 \pm 0.12 \mu m$ wide (n=30), and the PE2 cells were on average $1.9 \pm 0.65 \mu m$ long and $0.54 \pm 0.18 \mu m$ wide (n=30). These values were not significantly different from each other. Some cells were significantly longer than the average and were not measured on the basis they were undergoing cell division. On some occasions preparations of the wild type and insertional knockout mutant cells displayed straight morphologies. However no coccoid cells were observed.

A similar EM analysis of *C. jejuni* 81-176 cells revealed that these cells also displayed spiral morphology with bipolar flagella (Figure 3.11). The PE4 cells were on average $2.26 \pm 0.77 \mu m$ long and $0.58 \pm 0.20 \mu m$ wide (n=30), and the PE5 cells were on average $2.01 \pm 0.85 \mu m$ long and $0.54 \pm 0.15 \mu m$ wide (n=30). Although there was a small difference in the length of PE4 and PE5 this was not significant. As with 11168H some cells were significantly longer than the average and were not measured on the basis they were undergoing cell division. For both strains the loss of the P450-encoding gene does not affect the overall morphology of the cells when viewed by EM. It has been reported that the disruption to the *cj1411c* gene results in “fattening” of *C. jejuni* 11168 cells [202]. However, this phenotype was not observed in these analyses.
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

Figure 3.9. Transmission electron micrographs investigating the effects of *cj1411c* insertional knockout in *C. jejuni* PE1 and PE2 strains. Both PE1 and PE2 strains exhibited spiral morphology and bipolar flagella. Scale bar lengths are shown.
Figure 3.10. Transmission electron micrographs investigating the effects of *cj1411c* insertional knockout in *C. jejuni* 81-176 PE4 and PE5 cell strains. Both PE4 and PE5 strains exhibited spiral morphology and flagella. Scale bar lengths are shown.
3.8. The role of CYP172A1 in flagella-mediated motility

The results from AAG an EM experiments demonstrated that flagella were present (Figures 3.8 to 3.11). This implied that gross assembly was unlikely affected by the loss of CYP172A1, but the role of cj1411c in flagella-mediated motility had not been studied. In order to investigate this further, C. jejuni cells were harvested, standardised to $10^5$ cfu ml$^{-1}$, and 10 µl inoculated into 0.4 % MH agar (Section 2.13). This reduced percentage agar allows the bacteria to use flagella-mediated motility to migrate through the agar.

There was no difference in motility between PE1 and PE2, and between PE4 and PE5 strains (Figure 3.12). After 48 hours there is no difference between the wild type and CYP172A1 lacking mutant cells, where the 11186H strains produced a diameter of motility of 10 mm compared to 12 mm in the 81-176 strains. Though there appeared to be a slight difference in motilities after 24 hours, this was not statistically significant. During the assays, both the wild type and the insertional knockout mutant strains produced chemotactic rings, which have been observed previously in other C. jejuni motility experiments [203, 204].
Figure 3.11. Effect of *cj1411c* insertional knockout on flagella-mediated motility of *C. jejuni* 11168H and 81-176 strains. PE1, PE2, PE4 and PE5 strains were point inoculated into 0.4% MH agar and incubated for 48 hours. Diameters of motility were measured after 24 and 48 hours. Error bars indicate SE, n=3.
3.9. The role of CYP172A1 in determining the sensitivity of \textit{C. jejuni} to antimicrobial agents

MIC assays, using the standardised broth dilution method (Section 2.14) [205], were performed to investigate if the loss of CYP172A1 altered susceptibility to a range of antimicrobials affecting DNA replication, protein synthesis, peptidoglycan synthesis and cell membrane integrity. Growth inhibition was defined as a non turbid microwell with no difference in absorbance compared to control wells.

The results from three independent experiments indicated that the susceptibilities of \textit{C. jejuni} 11168 and 81-176 strains to the different antimicrobials were very similar (Table 3.2). The wild type (PE1 and PE4) and insertional knockout mutant (PE2 and PE5) strains displayed identical MIC values against all but three compounds (Table 3.2). The majority of the MIC values obtained in this study were generally consistent with previously reported values [179, 206], although MIC values for β-lactams ampicillin (16 µg/ml) and cefotaxime (8 or 16 µg/ml) were higher.

There were two antimicrobial agents, polymyxin B and sodium deoxycholate that produced a consistent difference in MIC between the wild type and insertional knockout mutant strains. Sodium deoxycholate is a water soluble form of bile acid and a detergent. Detergents emulsify cell membranes causing lysis and cell death [207]. Polymyxin B specifically targets Gram negative bacteria and, also disrupts the cell membrane. It binds to lipids, altering membrane structure, and increases permeability allowing the influx of water leading to cell death [208]. \textit{C. jejuni} PE5 also demonstrated an increased sensitivity towards SDS compared to the PE4 strain. SDS is a detergent with a 12 carbon chain length and disrupts the integrity of the lipid membrane. SDS also has a polar sulphate group which can bind to polar proteins and interact with the solvent.
### Table 3.2. The role of CYP172A1 in conferring antimicrobial sensitivity in *C. jejuni* 11168H and 81-176 strains.

Cells were added to a serial dilution of the listed antimicrobials in a 96 well set up. Cells were incubated for 24 hours. The MIC value was calculated as the lowest concentration of the antimicrobial that prevented cell growth. *n=3.*

<table>
<thead>
<tr>
<th>Compound</th>
<th>PE1</th>
<th>PE2</th>
<th>PE4</th>
<th>PE5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Cefotaxime</td>
<td>8</td>
<td>8</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Gentamycin</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nalidixic acid</td>
<td>16</td>
<td>16</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Polymyxin B</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>SDS</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>256</td>
</tr>
<tr>
<td>Sodium deoxycholate</td>
<td>25,000</td>
<td>12,500</td>
<td>25,000</td>
<td>12,500</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>&lt;0.125</td>
<td>&lt;0.125</td>
<td>&lt;0.125</td>
<td>&lt;0.125</td>
</tr>
</tbody>
</table>
3.10.1. The role of CYP172A1 in sensitivity to polymyxin B, using E-Test strips
The increase in sensitivity of the *cj1411c* insertional knockout mutant to polymyxin B, compared to the corresponding wild type strain, was further investigated using polymyxin B E-Test strips (bioMérieux). Strips were aseptically transferred onto non selective CBA which had been freshly inoculated with either wild type or insertional knockout strains and incubated for 48 hours under standard conditions. The MIC value was then recorded as set out by the manufacturer's protocol.

A consistent difference in sensitivity was observed between PE1 and PE2 strains, with the former having an MIC of 16 µg/ml and the latter 12 µg/ml. The increase in susceptibility was reversed with *cj1411c* gene complementation (PE3 strain), and restoration of the MIC to wild type levels (16 µg/ml) was observed. The results from the 81-176 strains were similar. PE4 had an MIC value of 16 µg/ml and PE5 was more sensitive to polymyxin B (MIC 8 µg/ml). However, the reintroduction of the *cj1411c* gene (PE6 strain) resulted in only partial restoration of polymyxin B sensitivity to wild type levels (MIC value 12 µg/ml). The increased sensitivity to polymyxin B is not as dramatic as observed with the severe truncation of LOS caused by mutation of the *waaF* gene (responsible for attachment of the second heptose in the LOS core, see Figure 1.7). Disruption to the *waaF* gene in *C. jejuni* 81-176 results in a 6 fold increase in polymyxin B susceptibility, which was partially restored by complementation of the gene [209].

3.10.2. The role of CYP172A1 in sensitivity to polymyxin B, using agar dilution
The agar dilution method was performed (to National Committee for Clinical Laboratory Standards guidelines) as previous studies have shown that E-Test strip and agar dilution can be complementary methods for determining MIC values [210, 211]. Agar dilution involves mixing a concentrated stock of polymyxin B solution into molten CBA before dispensing into petri dishes. Cell suspensions (10 µl) were then applied to these plates and viability calculated similar to the Miles and Misra technique (Appendix Figure S3). Control plates with no polymyxin B added were used to standardise the results so that direct comparisons from multiple experiments could be achieved.
A significant difference in the polymyxin B sensitivity between the wild type strains (PE1 and PE4) and the insertional knockout mutant strains (PE2 and PE5) (Figure 3.13) was observed. PE2 was unable to grow on agar that contained more than 8 µg/ml polymyxin B. The difference between the survival rates of the PE1 (78 %) and PE2 (22.5 %) at 8 µg/ml was statistically significant (p ≤0.05). PE5 also displayed an increase in susceptibility compared to the PE4 and was unable to grow on agar containing more than 10 µg/ml polymyxin B, whereas PE4 was able to grow on agar containing 16 µg/ml polymyxin B. The difference between the survival rates of the PE4 (88.7 %) and PE5 (0.6 %) at 8 µg/ml and 72 % and 0.04 % at 10 µg/ml polymyxin B were statistically significant (p ≤0.05).

The results, for both 11168H and 81-176, indicated that the loss of cj1411c gene product increased polymyxin B sensitivity, and this was reversed by genetic complementation. This might suggest that an alteration to the bacterium surface results in an increase polymyxin B accessibility. The MIC values obtained using the agar dilution method were very similar to those recorded using the E-test strips. This further validated and confirmed that the P450 lacking mutant was more susceptible to polymyxin B than the wild type strains.
Figure 3.12. The role of CYP172A1 in mediating polymyxin B sensitivity in *C. jejuni* 11168H and 81-176 strains to polymyxin B containing agar plates. Cell survival was compared to bacteria not exposed to polymyxin B. A) Comparison of polymyxin B sensitivity between PE1 (blue) and PE2 (red) strains. PE2 displayed an increase in susceptibility compared to PE1 and was unable to grow on agar with >10 µg/ml polymyxin B. B) Comparison of polymyxin B sensitivity between PE4 (blue) and PE5 (red) strains. PE5 displayed an increase in susceptibility compared to PE4 and was unable to grow on agar containing >10 µg/ml polymyxin B. Error bars = SE, n=3 * indicates significant difference between wild type and P450 lacking strains using a student’s paired T-test (p<0.05).
3.11. The role of CYP172A1 in the sensitivity of C. jejuni to antifungal azole compounds

Azoles are potent antifungal agents and their primary mode of action is to inhibit the P450 (CYP51) mediated 14α-demethylation of lanosterol [100]. The sensitivity of C. jejuni strains to a range of azoles (50 mg/ml), including econazole, fluconazole and miconazole, was investigated to establish if CYP172A1 might be implicated (like CYP51) in the maintenance of the fluidity of the cell membrane. Although azoles did affect C. jejuni growth, there were no observable differences in sensitivity between the PE1 and PE2 cells for any of the tested azoles (Figure 3.13). In both strains, econazole (disc 1) was the most potent growth inhibitor followed by miconazole (disc 3) and ketoconazole (disc 4). Clotrimazole (disc 2) and voriconazole (disc 6) failed to inhibit growth more than that observed for the DMSO control (the solvent used to dissolve the azole drugs). These data suggests that econazole (and other azoles) cause bacteriostasis via a P450 independent mode.

The sensitivity of PE1 and PE2 to econazole was further investigated to ensure that there was no difference in inhibition of cell growth over a range of different concentrations (50, 10, 2 and 0.2 mg/ml) (Figure 3.14). Both strains had similar sensitivities in the dilution range, whereby 50 to 2 mg/ml econazole solutions were effective in inhibiting growth, but at 0.2 mg/ml no inhibition above the control was observed. The results further suggest that growth inhibition occurs via a P450-independent mechanism. Recently, it was discovered in Candida albicans that the uptake of azoles is not a passive process, but instead occurs via facilitated diffusion [172]. The mechanism and proteins involved in this process are as yet unknown, but it might be possible that Campylobacter also has a similar type of transport mechanism, and that this enables the accumulation of econazole, which may be bacteriostatic or otherwise lead to bacterial cell death.
Figure 3.13. The role of CYP172A1 in mediating azole sensitivity of *C. jejuni* PE1 and PE2 strains. A disc diffusion assay indicating the sensitivity of PE1 (A) and PE2 (B) cells to azole drugs. A range of azoles were tested (all at 50 mg/ml): econazole (1), clotrimazole (2), miconazole (3), ketoconazole (4) and voriconazole (5). A disc containing DMSO (D) was used as a control. Econazole is clearly the most effective azole drug in this assay.

Figure 3.14. Concentration dependence of econazole mediated bacteriostasis in *C. jejuni* PE1 and PE2. A disc diffusion assay of econazole with PE1 (A) and PE2 (B) cells. A range of econazole concentrations were tested: 50 mg/ml (1), 10 mg/ml (2), 2 mg/ml (3) and 0.2 mg/ml (4). A disc containing DMSO (D) was used as a control. Zones of inhibition demonstrate clearly the inability of cells to grow in the presence of the indicated concentrations of econazole.
3.12. The role of CYP172A1 in LOS biosynthesis in \textit{C. jejuni}

The LOS is an important structure on the cell surface, and although it might be assumed that CYP172A1 has a role in CPS biosynthesis it was still important to examine the effects to the LOS upon the loss of CYP172A1 function. The LOS of \textit{C. jejuni} is comprised of a lipid A moiety with core oligosaccharide attached [60] and is readily stainable with silver ions. Silver staining is a very sensitive procedure that has a detection limit lower than that of coomassie stain [212]. Cell samples were standardized to an OD$_{600}$ of 1 in SDS buffer, treated with proteinase K to digest proteins and electrophoretically separated by SDS-PAGE [213] (Section 2.16).

The mobility of the LOS species from the PE1 and PE2 cells of \textit{C. jejuni} 11168H were identical with an apparent M$_w$ of $\sim$10 kDa (Figure 3.16). The technique is able to resolve differences in structure of a single sugar from the oligosaccharide. The \textit{Acj1139::kan} mutant carries an insertional mutation within a β-1,3 galactosyltransferase, a phase variable gene that adds a single galactose residue to the terminal GalNAc on the LOS core [59]. This mutant was selected as a control in case either PE1 or PE2 strains displayed phase variability of this gene, which could be misinterpreted as being a significant result. The result also indicates that the \textit{cj1139} gene was expressed in 11168H as the mobility of the LOS in PE1 and PE2 strains was less when compared to the \textit{Acj1139} mutant. A similar result was obtained for the 81-176 PE4 and PE5 strains, whereby there was no observed difference in the mobility of the LOS species. From these results it appeared as though there was no alteration to the sugar composition of the LOS in the \textit{cj1411c} insertional knockout mutant strains.
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

Figure 3.15. Effect of *cj1411c* insertional knockout and complementation on LOS mobility in *C. jejuni* 11168H. Standardised samples of *C. jejuni* PE1, PE2 and PE3 and Δcj1139::kan mutant strains were loaded onto and resolved on a 12 % SDS-PAGE gel. The gel was then stained using silver stain. The LOS of the samples was estimated to be just over 10 kDa except for the Δcj1139 mutant, which was estimated to be ~10 kDa.
Chapter 3. Investigating the role of the cj1411c product, CYP172A1, in Campylobacter jejuni

3.13. The role of CYP172A1 in the biosynthesis of CPS in C. jejuni

Alcian Blue can be used to stain C. jejuni CPS in polyacrylamide gels, and on the surface of intact cells [77, 81]. To visualise CPS whole cell lysates were resolved using tricine PAGE gels. Gels were fixed and stained with a 0.5 % solution of Alcian Blue as described in the methods (Section 2.16) [66].

C. jejuni PE1-PE5 cell samples from were analysed along with Δcj1139, and capsular deficient mutant (ΔkpsM [66]) strains (Figure 3.17). Only the LOS band was detected in all the samples after staining, with a M\textsubscript{w} of ~10 kDa, except for the Δcj1139 which displayed an increased mobility. A broad band would be expected between 15-20 kDa (except for the kpsM lacking mutant) in successful CPS staining, but in all the samples this was not detected [66]. The experiment was repeated several times with the same outcome, and even with additional cellular material loaded, no CPS band was detected. In order to determine that the lack of a visible band was not a result of the growth conditions the experiment was repeated using cells grown in a liquid media and at 37 °C. In all cases no CPS was detected, but LOS was visualised.

To further investigate the poor binding of alcian dye to the C. jejuni 11168H and 81-176 CPS, a range of other C. jejuni strains were tested for their ability to bind the dye. Insertional knockout mutants of the kpsM gene were available for C. jejuni P and G and their capsules had been previously detected used Alcian Blue dye [66]. The samples were prepared as before (grown on CBA at 42°C under microaerobic conditions). A band with the mobility consistent with LOS was detected in all strains (Figure 3.18). In all the samples, except P2 kpsM mutant (P2.a4), a second band was visible at ~20 kDa, the identity of which is unknown. However, the band was not broad, suggesting that it was not a CPS signal, coupled to the fact that the band is detected in both the wild type and kspM mutants, indicates that the band is possibly a non degraded contaminant common to all the strains.

Difficulty in visualising CPS from C. jejuni 11168H has been reported, and is possibly due to the lack of charge on the CPS for the dye to bind to [77]. Despite the use of a wide variety of conditions and using a broad range of C. jejuni strains the visualisation of the capsule was not achieved using Alcian Blue. The stain was able to detect the LOS
of the cell lysates but no broad band ~15-25 kDa was not detected, which would be indicative of CPS [66, 214]. Due to the inability to visualise the CPS via this method, spectroscopic techniques were employed to discover if the capsule was present or absent in the insertional knockout mutant.
Chapter 3. Investigating the role of the cj1411c product, CYP172A1, in Campylobacter jejuni

Figure 3.16. Effect of cj1411c insertional knockout and complementation on CPS structures in C. jejuni 11168H and 81-176. Alcian Blue staining of resolved standardised cell samples was used in an attempt to visualise the LOS and CPS structures. Standardised cell samples of PE1, PE2, PE3, ΔkpsM and Δ1139 mutant strains from 11168H and PE4 and PE5 from 81-176 strains from liquid culture were resolved on a 12.5 % tricine gel. A LOS band was detected in all samples just over 10 kDa except for the Δ1139 which ran to 10 kDa.

Figure 3.17. Alcian Blue staining of C. jejuni 11168H P and G strains and their respective ΔkpsM mutants. Samples, grown on solid phase media, were resolved on a 12.5 % tricine gel. The gel was then stained using Alcian Blue dye. All of the samples run at ~10 kDa. A second band was observed ~20 kDa and represents an undigested contaminant. The “a4” denotes a ΔkpsM mutant.
3.14. Fourier Transform Infrared (FT-IR) spectroscopy of Campylobacter species

FT-IR works by measuring the vibrational properties of chemical bonds excited by infrared radiation. Different types of bonds produce different vibrations and in a biological sample the wavelengths at which the vibrations occur can be assigned to fatty acids (3000-2800 cm\(^{-1}\)), amides/proteins (1750-1500 cm\(^{-1}\)), mixed carboxylic groups (1500-1250 cm\(^{-1}\)), RNA/DNA (1250-1200 cm\(^{-1}\)) and polysaccharide (1200-900 cm\(^{-1}\)) [182, 215].

Over the past 20 years there have been a number of publications relating to the use of FT-IR as a tool for identifying microorganisms including E. coli, Bacillus and Campylobacter [216-218]. There are many benefits in using FT-IR for microbial identification. It is rapid, non-damaging, large data sets can be collected for better discrimination of closely related organisms and organisms that are hard to identify using classical techniques using biochemical and physiological properties can be analysed [219]. Due to these properties FT-IR was chosen as a method of profiling Campylobacter species and strains with the purpose of identifying a function of the \(cj1411c\) gene.

A range of different fast growing Campylobacter species (Section 2.4) were cultured for 24 hours under microaerobic conditions. No antibiotics were used during the growth phase in order to limit the number of variables that might affect the differences observed between the different species. Cell samples were then measured by FT-IR and analysed as described (Section 2.18.1). For each cell type three biological replicas and three sample replicas were measured, producing nine different spectra for each cell type. The increase in sample acquisition enhances the reliability of the data acquired [220]. FT-IR was used to discriminate between the different Campylobacter species. Initially, the raw spectra from the averaged data acquisition for each species were analysed, but for all the species the spectra were very similar (Figure 3.19). Major peaks were observed in the same regions as would be predicted for a biological cell sample (e.g. amide, DNA and polysaccharide). From visual comparison of the different spectra, it was not possible to determine one species from another.
Pattern recognition was then implemented, to ascertain if any relationships and/or differences between the species was possible. This was achieved using supervised discrimination function analysis (DFA), in which *a priori* knowledge of which spectrum relate to which sample is provided. This can help to reduce the impacts of spectrum variation between related samples that is caused by non-sample signals. By performing DFA the dimensionality of the data is reduced into principal components (PCs), with much of original data variability described with just a few linear combinations, which allowed simpler data analysis and visualisation. Ordination plots using the first two PC scores allowed for the best representation of the variance associated with all the spectra [183, 221]. When using DFA it is important to test the model that you create, which involves the evaluation of how well the model can predict which “class” a spectrum belongs to when it is not identified. For this work, Fisher’s linear discrimination analysis was used and the strength of the model scored from 0-100 %, where a score of 100 % corresponds to a model, which accurately identifies an unknown spectrum into the correct “class” 100 % of the time [221].

DFA analysis of the spectra collected from the different *Campylobacter* species resulted in the separation of all the samples (Figure 3.20). The first two discrimination factors (principal components) were able to separate out 73.6 % of the data. Closely related *C. jejuni* strains clustered together in the top right quadrant of the ordination plot (blue circle). With the exception of *C. jejuni* 11168H and 81-176 the strains formed clear and separate groups. More distantly related species formed their own groups that were clearly identifiable from other species. The accuracy of the model used to discriminate the spectra was 100 % in all samples, with the exception of the *C. jejuni* 81-176 sample, which scored 78-81 % accuracy. This score is reflected in the ordination plot, where the 81-176 strain partially overlapped with the 11168H strain. The separation of *Campylobacter* species has previously been demonstrated with confocal Raman spectroscopy [222] and the results presented illustrate that the same level of separation is achieved using DFA analysis from data collected by FT-IR.
Figure 3.18. Averaged FT-IR spectra of *Campylobacter* species and *C. jejuni* strains. Spectra from all the samples were overlapped to allow for easier comparisons. All spectra produced typical signals for cellular components, including fatty acid (3000-2800 cm$^{-1}$), amides/proteins (1750-1500 cm$^{-1}$), mixed carboxylic groups (1500-1250 cm$^{-1}$), RNA/DNA (1250-1200 cm$^{-1}$) and polysaccharide (1200-900 cm$^{-1}$).

Figure 3.19. PC-DFA ordination plot of *Campylobacter* species and *C. jejuni* strains. The *C. jejuni* strains group together in the top right quadrant of the plot. Data were collected in quintuplet from three biological replicas. Two PC’s were used to create the plot, covering 73.6% of the explained variance. The accuracy of the model for DFA analysis is displayed in the legend.
3.15. Discrimination of *C. jejuni* PE1-PE6 strains by FT-IR spectroscopy

Following the discrimination of *Campylobacter* species, the next step was to investigate if the separation of *C. jejuni* PE1-PE6 strains was possible. Discrimination of different strains of *E. coli* has previously been reported [183]. However, in the current work discrimination at a sub-strain level was being investigated. The capsule deficient mutant Δ*KpsM* was included as a control to examine if CPS loss occurred in PE2 and PE5 strains. Cell samples were grown on non selective plates for 24 hours before being harvested and prepared for FT-IR as previously described.

The averaged spectra from all the samples were examined and appeared to be identical on visual inspection (Figure 3.21). This result was not unexpected based on the fundamental similarities between all the samples. As no differences were observed, DFA discrimination was performed.

Each of the three biological replicas were analysed separately in order to examine the biological variation between them (Figure 3.22). The first two components of discrimination analysis were able to separate out ~95 % of the data which shows very good separation of the data. The biological replicas produced similar results relative to each other, indicating the reproducibility of the data (a recognised feature of FT-IR). *C. jejuni* 11168H Δ*KpsM* mutant separated from the other 11168H strains which may be suggestive that the CYP172A1 mutant (PE2) has not lost its capsule coat. Analysis of PE1-PE3 strains was promising, whereby PE2 is almost forming its own distinct group and PE1 and PE3 strains overlap each other. The level of group separation was much greater in 81-176 strains (PE4-PE6) where all three cell types form very clear groups away from each other. With the exception of PE6, the two strains (11168H and 81-176) were separable, which was an improvement from the comparison of the FT-IR spectra in the larger sample data set (Figure 3.20.), and illustrates the resolving power of this technique. The accuracy of the prediction model was 100 % for PE4, PE5 and Δ*KpsM*. For the other samples the accuracy was lower, due to overlap on the ordination plot. PE3 (72-79 %), PE6 (71-78 %), PE2 (65-71 %) and PE1 (50-58 %).
Figure 3.20. Averaged FT-IR spectra of C. jejuni strains. Averaged spectra were overlapped on a single trace for comparison. All spectra produced typical signals for cellular components, including Fatty acid (3000-2800 cm$^{-1}$), amides/proteins (1750-1500 cm$^{-1}$), mixed carboxylic groups (1500-1250 cm$^{-1}$), RNA/DNA (1250-1200 cm$^{-1}$) and polysaccharide (1200-900 cm$^{-1}$). 81176 wt (PE4), 81176 ko (PE5), 81176 comp (PE6), 11168 wt (PE1), 11168 ko (PE2) and 11168 comp (PE3).
Figure 3.21. PC-DFA ordination plots of three biological replicas of *C. jejuni* strains. Data were collected in triplicate from three samples, totalling 9 spectra for each strain from each plate. Two PCs were used for PC-DFA which accounted for ~95% of explained variance. 81176 wt (PE4), 81176 ko (PE5), 81176 comp (PE6), 11168 wt (PE1), 11168 ko (PE2) and 11168 comp (PE3).
3.16. Probing the surface of *C. jejuni* using Surface Enhanced Raman Spectroscopy (SERS)

SERS is a technique that can be used to look at the surface structure of many materials including cells. This is in contrast to FT-IR which is able to measure the contents of an entire cell. The addition of a roughened metal, typically silver or gold, enables an increase in Raman scattering and improves the signal to noise ratio, making detection limits much lower, and improves the sensitivity to allow for detection down to single molecule level [223]. Because of this sensitivity it is vitally important that sample preparation is kept as uniform as possible to obtain precise results. SERS has successfully discriminated between different classes of bacteria, for example *Enterococcus sp.* and *E. coli*. It has also successfully discriminated different *E. coli* strains [184]. There are some cautions to observe when conducting SERS analysis beyond the consistency of sample preparation. The use of the laser in the presence of a colloid solution is known to irreversibly damage cells, which will drastically alter the spectra [224]. In order to overcome this, the position of the laser can be moved to another group of cell. However this then contributes to a new issue. The sample preparation is inhomogeneous and it is certain that different locations will produce different spectra, especially with regards to the distance between the cell and the colloid solution [224]. The collection of a large data set and the removal of outlying data can be used to reduce the impact of inhomogeneity.

Unlike FT-IR, there are currently no databases for the spectral assignments, which is not helped by different groups reporting different spectral features for identical bacterium [224]. This means that any spectral features need to be validated with control samples. One report has assigned raman wavelengths to a range of different chemical species of *C. jejuni*, including phenylalanine, amides and glycosidic rings, however their experimental design was different and so these values could not be used in this study [225].

*C. jejuni* strains PE1-PE6 and a ΔkspM mutant (from *C. jejuni* 11168) were investigated by SERS, using experimental conditions similar to those used for the FT-IR analyses (see section 2.18.2). Random locations on the CaF$_2$ disc were measured and 50 spectra
were acquired for each cell sample in an attempt to generate a large dataset, in order to improve the signal to noise ratio. The spectra collected were baseline corrected to remove any interference from high fluorescent background and all the spectra were normalised to allow direct comparison. A robust PCA was then applied to the data in order to identify any major outlying data (roughly half of the collected spectra), which were subsequently removed from the data set. The outliers could be attributed to spectra acquired in the absence of silver colloid, a blank spectra or spectra obtained from lysed cells [224]. The remaining spectra were then analysed by PCA (Figure 3.23.).

The results displayed on the ordination plot did not produce any definitive results. The CaF$_2$ discs (black circle) separated from the other data, which proved that the SERS was able to discriminate between discs with and without cells present. The TEV from the 2 main principle components was just over half of the entire variance in the samples (54.1 %). This value was lower than observed for the FT-IR and might have been caused by a lower than expected signal to noise ratio which would have affected the PCA scoring. Each of the strains was identifiable as its own cluster, with the exception of PE2 and PE3 (yellow and orange, respectively), which suggested that SERS had the ability to separate C. jejuni at a subspecies level, a greater resolving power than previously obtained for E. coli [184]. PE6 (pale blue) formed a cluster far from the other cell samples. Analysis of the raw spectra revealed that the reason for this was the lack of a signal below 1000 cm$^{-1}$ and a large signal at 2000 cm$^{-1}$ which were not present in the other cell spectra (Appendix, Figure S2). The identity of this peak is unknown and might be due to a mild contaminant that produced a strong interaction with the silver colloid.
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

Figure 3.22. PCA ordination plot showing clusters obtained from the analysis of selected SERS data obtained in *C. jejuni* strains. PCA obtained from SERS analysis of *C. jejuni* strains PE4 (cell 1), PE5 (cell 2), PE6 (cell 3), PE1 (cell 4), PE2 (cell 5), PE3 (cell 6) and ΔkspM (cell 7). A blank disc was also analysed (cell blk). PC: principal component, TEV: total explained variance.
The results were not as expected. In theory the wild type and complement strains would have overlapped (PE1 to PE3 and PE4 to PE6) with the P450 insertional mutant possibly forming its own cluster. The 11168 ΔkpsM mutant (red) did form its own cluster away from the other 11168H strains, suggesting that loss complete loss of the capsular coat had not occurred in the P450 lacking mutant. However, the location of PE2 and PE3 are nearer to the capsular deficient mutant and so it cannot be ruled out that an alteration to the capsule has occurred. SERS analysis of *C. jejuni* is a new field with only a single report describing attempts to characterise spectral features observed from this bacterium. An improvement to the discrimination power of SERS might be achieved with the identification of “SERS-active hot-spots”, areas where the position of the silver colloid is optimal for enhancement of cell surface components [226]. The use of SEM would allow for the selection of optimum sampling, which could be used to identify cells is the same orientation, relative to each other, potentially removing a large source of experimental error.
3.17. Summary

The *cj1411c* gene of *C. jejuni* was successfully knocked out, and then complemented in *C. jejuni* 11168H and 81-176, as confirmed by PCR analysis. It was important to ensure that the orientation of the inserted kanamycin resistance cassette was retained following any genetic manipulations performed, in order to prevent polar effects, including effects on the transcription of genes downstream from *cj1411c*, which would affect any interpretations made from the comparative data collected on the properties from PE1-PE6 strains. The creation of the insertional knockout mutant and complement strains allowed for the exploration of the role of the *C. jejuni* P450 (CYP172A1). A range of techniques were employed to elucidate the function of this enzyme.

Importantly, the production of CYP172A1 in *C. jejuni* was demonstrated for the first time by western blotting in the model strains 11168H and 81-176. This builds on the work by Parkhill *et al.*, who initially identified the *cj1411c* gene in the *C. jejuni* genome, and later work which identified that mRNA from the *cj1411c* gene in the 11168 strain and *C. lari*, indicating that the gene is expressed [14, 190, 227]. CYP172A1 was detected in the wild type (PE1 and PE4) and complement (PE3 and PE6) strains, but not in the insertional knockout strains (PE2 and PE5). Initially, the identification of CYP172A1 proved difficult due to the large amount of cross reactivity observed with antibodies also targeting other *C. jejuni* proteins. These difficulties were reduced by mixing the antiserum with PE2 cell lysate and by also utilising the recombinant CYP172A1 as a positive control. The antiserum allowed for the detection of CYP172A1 in wild type and complemented cells only, albeit with the levels observed in the complement being slightly lower compared to the wild type stains (Figure 3.6), and that the insertional knockout resulted in the loss of P450 production. The presence of the enzyme in cells grown at both 37 °C and 42 °C is an interesting result, and would suggest that the P450 is constitutively expressed at both these temperatures and is unlikely to exhibit temperature dependent expression, a phenomenon which is observed for a number of other proteins in *C. jejuni* (e.g. PEB1, HisD and Tpx) [193, 194]. This might indicate that the P450 is not directly involved in human pathogenesis, but performs a housekeeping function within the cell, possibly having a metabolic or biosynthetic role.
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

The P450 gene (*cj1411c*) is not essential to the bacterium’s survival, as the *cj1411c* knockout mutants created were still viable. A discrepancy between optical density reached by wild type and insertional knockout cultures and the viable cell counts was observed. Both the wild type and insertional knockout strains had very similar viable counts, as determined by the Miles and Misra technique. However, the optical densities (OD$_{600}$) of the *cj1411c* insertional knockout mutants were lower than the wild type strains throughout the time course of the growth assay. This characteristic has not been previously reported. A possible explanation involves how optical density measurements are recorded. A spectrophotometer emits an incident light that passes through a sample cell and the light then hits the detector. For bacterial growth assays, a common misconception is that the spectrophotometer measures the light that is absorbed by the sample. However, the measurement made for microbial cell growth at 600 nm instead relates to the amount of light scattered by the turbid cell suspension [228]. A more turbid sample will scatter more light than a less turbid sample, producing a higher apparent absorbance reading. If there is a difference in the cellular structure between the wild type and insertional knockout strains that affects the light scattering, then this may explain the difference in the observed OD$_{600}$ readings, despite these cultures displaying essentially the same viable counts.

Results from the AAG and cell motility experiments suggested that there might be a surface alteration to the bacterium when the *cj1411c* gene is knocked out. The AAG indicated that the insertional knockout mutant had a decreased AAG when compared to the wild type and complement strains in both 11168H and 81-176 strains. However, in the case of 11168H, when compared to a mutant lacking flagellin A, the effect of the loss of CYP172A1 does not seem to be as disruptive to the AAG in comparison to the loss of flagella. AAG is affected by the flagella and by surface charge [54, 175, 199]. As the P450 mutant AAG was not as severely affected as that in the flagella mutant, the result may be suggestive of some other type of surface-related change, and possibly one which involves alterations to the bacterial surface charge. The motility assay data support some of the conclusions drawn from the AAG data. The PE1, PE2, PE4 and PE5 strains displayed no measurable difference in motility, and so it might be concluded there is no significant change in the structure, the glycosylation levels or the functional
capability of the flagella. However, investigations into the flagellar structure and glycosylation levels would be required to confirm this hypothesis.

There were no obvious differences observed between the wild type and \textit{cj1411c} insertional knockout strains with respect to their overall morphology. This result is in disagreement with a previously published report, which suggested that the loss of the P450 function caused the cells to become shorter and fatter [202]. The strains used in the present studies were prepared on several occasions, but this short/fat phenotype was never observed. For some preparations, the classical S shaped morphology was lost, and the cells appeared as straight rods. It is possible that these cells were in the process of shifting to a coccoidal form prior to the blebbing stage, as previously reported [195].

The effects of a range of antimicrobial agents on the wild type and \textit{cj1411c} insertional knockout strains of \textit{C. jejuni} were investigated. The antimicrobials used have different modes of action (e.g. affecting protein synthesis, DNA replication \textit{etc}). The CYP172A1 insertional knockout mutants displayed no significant change in sensitivity to the agents used, with the exception of the detergent-like molecules. The site of action for these detergents is primarily at the outer membrane of the bacterium [229]. They are able to interact with the membrane and disrupt the stability of the lipids, which are important in maintaining the cell’s structural integrity. Polymyxin B is a detergent that is well known to specifically target gram negative bacterial membranes (including those of \textit{C. jejuni}) [230]. Further experiments investigating the sensitivity towards polymyxin B were conducted using E-Test strips and agar dilution methods. An increased sensitivity occurred with the loss of the \textit{cj1411c} gene, but the sensitivity was restored with the reintroduction of the gene in the complement strains. Polymyxin B sensitivity was shown to be linked to the integrity of \textit{C. jejuni} LOS, where by the loss of the LOS results in an increased polymyxin B sensitivity [209, 231]. \textit{C. jejuni} can modify the LOS to reduce the binding of polymyxin B and other cationic antimicrobials by decreasing the negative charge associated with the LOS by adding other polar groups (e.g. phosphoethanolamine, 4-amino-4-de-oxy-L-arabinose, and/or palmitate), to the lipid A [232]. It is unlikely that the complete loss of LOS has occurred, as the change in sensitivity was not as large as previously observed in LOS-associated genes. It might be
possible that the loss of CYP172A1 increases the accessibility of polymyxin B to the outer lipid membrane, possibly by altering the structure or charge of the CPS.

The antimicrobial action of a group of antifungal compounds, the azoles, was also tested to see if they might provide a novel series of compounds capable of killing *C. jejuni*. Azoles have been successfully used to mediate the killing of *M. smegmatis*, at sub micro-molar concentrations, and econazole was demonstrated to clear *M. tuberculosis* from infected mice [100, 233]. Some of the azoles tested in these studies were able to prevent growth of *C. jejuni*. However, it remains unclear whether the azoles that were effective were bacteriostatic or bacteriocidal. Regardless, the results obtained indicated that the inhibitor action of the azoles was likely CYP172A1 independent, since there was no obvious difference between azole sensitivities of all the *C. jejuni* strains tested. The concentrations at which the most successful azoles (econazole, ketoconazole and miconazole) were effective were too high to have a clinical use, particularly given the fact they would be used systematically to treat *C. jejuni* infection and would likely cross-react with and inhibit human P450 enzymes. The azoles used were all highly hydrophobic and were all prepared in DMSO. It is thus not fully clear whether the active concentration of these drugs in the agar is inaccurate, since precipitation of some of the azoles may have occurred in the agar. Thus, it is possible their potency against *C. jejuni* is actually somewhat greater than the agar diffusion assays suggest.

The analysis of silver stained gels indicated that there was no major alteration to the LOS in the PE2 and PE5 strains. The use of the Δ1139 mutant was important in preventing an incorrect functional assignment to CYP172A1 that might have been observed due to the phase variation of the 1139 gene. The detection of *C. jejuni* CPS was not achieved using methods previously described [66, 77], and only the LOS was detected. Karylshev noted that the 11168 CPS did not interact as favourably with Alcian Blue as in other strains (e.g. *C. jejuni* G1 and 81116) probably due to different affinities of the dye towards the chemically different CPS structures. In order to investigate the effect of growth parameters on CPS production, PE1-PE3 strains were grown on solid and liquid media, and growth was conducted at both 37 °C and 42 °C. These parameters did not improve binding of the dye to the CPS. Finally, other *C. jejuni* strains which
have been previously stained with Alcian Blue dye (including G1) where tested for Alcian Blue binding. Once again the CPS was not visible in the strains tested.

As successful visualisation of the capsule was not achieved with Alcian Blue stain, alternative spectroscopic methods were used to investigate the differences between the *C. jejuni* strains PE1-PE6. FT-IR was used a potential rapid and accurate method for the typing of *Campylobacter* [222]. Raman spectroscopy is able to analyse and distinguish multiple components of the bacterium (e.g. capsule, lipid and DNA) simultaneously. However, there are several factors affecting the reproducibility of FT-IR, including growth temperature, incubation time, sample preparation and instrument variation, as well as *Campylobacter* shape. The older a culture becomes, and the more unfavourable the environment is, the more likely coccoid cells are to develop. As well as a morphological change, there are chemical changes associated with this process [234]. For this reason the age of the cultures used was kept to 24 hours in an attempt to minimise the potential variations between the cultures. The results obtained for the different *Campylobacter* species examined indicated that FT-IR was able to discriminate all the species tested. *C. jejuni* strains formed a collective group separate from the other *Campylobacter* species. A recent report that used a larger number of species also obtained similar level of species separation [222]. The PE1-PE6 strains were also analysed by FT-IR. Carbohydrates have defined vibrational patterns between 1200-900 cm\(^{-1}\), and so the profiles of the strains could be compared. Visual comparison of the spectra produced from each strain, and from the Δ*kspM* mutant did not reveal any distinguishing features and so discriminant functional analysis was performed. The first noticeable result from performing this analysis was that the 11168 Δ*kpsM* mutant and PE2 were not located at the same position on the ordination plot. This might suggest that the PE2 mutant is not capsule deficient. However, the discrimination could also be between other cellular components. Promisingly, the PE1 and PE3 strains did appear to overlap in three biological replicates and PE2 displayed signs of forming a separate group. However, the PE4, PE5 and PE6 strains formed three separate groups with PE6 being placed within the PE1-PE3 cluster, which was unexpected. Although separation of different genera and species has been achieved with varying levels of success, to date subspecies and sub-strain identification has not been achieved in *Campylobacter*. This is
Chapter 3. Investigating the role of the *cj1411c* product, CYP172A1, in *Campylobacter jejuni*

perhaps unsurprising, based on the strong biochemical similarities shared between these organisms. Thus, in attempting to resolve the PE1-PE6 strains spectroscopically, new territory was entered in this thesis. The data collected suggested that, while FT-IR technique has a potential in bacterial analysis and discrimination slight variations in bacterial composition may have major effects on the reliability and reproducibility of the results obtained.

In conjunction with FT-IR, an investigation into the surface properties of the *C. jejuni* strains was performed using surface enhanced Raman spectroscopy (SERS). Unlike FT-IR, SERS looks at only the surface properties of a sample. In the case of *C. jejuni* this involves mainly the CPS, LOS, flagellum and outer lipid membrane. The addition of a silver colloid allowed for an increased signal detection, improving the signal to noise ratio. Principal component analysis was conducted on the *C. jejuni* strains PE1-PE6 along with a capsular deficient mutant. This capsular mutant formed its own cluster in the ordination plot, suggesting that the full loss of the capsule (in strain PE2) had not occurred. The experiments produced unexpected results, whereby the complemented and wild type strains did not overlap in the ordination plot, as might have been expected. This might be explained by the nature of the sample used, which is inhomogeneous. On the CaF$_2$ disc, the orientation of the cells are not fixed, and nor is their distance from the silver colloid material. These two factors alone could have a large impact on the data acquisition and affect the averaged spectra collected for each strain, which could then influence the discrimination analysis on the samples.

The results from this chapter suggest that a change to the bacterium’s surface properties may occur with the loss of the *cj1411c* gene product. The reduced AAG and the decrease in optical density readings during growth indicate such a significant alteration. At present it cannot be determined what surface structure is responsible for this observed phenotype, as visualisation of the CPS was not achieved and so the role of CYP172A1 in CPS biosynthesis could not be adequately investigated. However, the data collected are suggestive that CYP172A1 is not directly involved in flagellum biosynthesis or function.
Chapter 4. Expression, purification and biophysical characterisation of CYP172A1

4.1. Introduction
Due to the recent discovery of the lone C. jejuni cytochrome P450, CYP172A1, in C. jejuni our understanding of the function of this protein, including its biophysical properties, is limited. CYP172A1 production was optimised in E. coli to allow purification of the protein in sufficient quantities to perform in vitro biophysical assays and to examine structure. Recombinant expression allows for the production of large amounts of soluble protein that can be easily purified, which would not be the case for the native protein from C. jejuni. Expression of non-native proteins using E. coli expression systems is a well recognised technique that has been applied in studies of multiple proteins. One of the many benefits is the ability to produce large quantities of protein using relatively inexpensive starting materials. This can allow for detailed experiments to be performed on expressed proteins, many of which may require large amounts of protein, particularly crystallography [235].

Although the primary sequence of CYP172A1 contains the highly conserved heme binding motif, FxxGxxxCxG, no previous studies had actually looked at confirming that the enzyme was indeed a P450. This was studied by examining the formation of a 450 nm Soret peak with the protein under reduced conditions and the solution saturated with CO. To date P450s are the only known proteins to give rise to this characteristic spectral peak. This was important as it would allow for further general P450 characterisations to be conducted (e.g. the examination of secondary and tertiary structures as measured by FUV CD and fluorescence spectroscopy respectively). It was also important to gain an appreciation for the structural features of the enzyme. To date there are 27 unique structures. However, the fact that CYP172A1 shows such low homology to any other structurally solved P450 suggests that there is a chance that the topology of this enzyme might be novel. A novel structure might also display different biophysical characteristics, such as stability and secondary structural organisation, and so it was also important to investigate these aspects further. The generation of protein crystals would
allow for the three dimensional structure of CYP172A1 to be determined, which is critical for understanding the topology of the active site, thus providing a way of determining the likely natural substrate. For this reason, crystallisation trials were also undertaken in efforts to solve the CYP172A1 structure.

This chapter thus aimed to confirm test whether or not CYP172A1 is indeed a cytochrome P450 and, by using a combination of biophysical and spectroscopic techniques, also probe the biochemical nature of this enzyme and see if it is comparable to other characterised P450s. The generation of protein crystals would allow for the three dimensional structure of CYP172A1 to be determined, and would allow for accurate predictions of the natural substrate to be made based upon the protein’s active site.

4.2. Construction of the CYP172A1 recombinant protein expression system

A codon optimised version of the cj1411c gene was created by GeneService and inserted into a pET15b vector (unpublished work by Dr Kirsty McLean, University of Manchester). The pET15b vector enables the attachment of a N-terminal His$_6$ tag to the P450 to allow for the partial purification of CYP172A1 from an E. coli expression cell lysate using Ni-containing resin. The vector also allowed for the expression of the P450 to be controlled by the T7 lac promoter in an IPTG-inducible manner. The PET15b vector also carries an ampicillin resistance cassette to allow for selective screening of bacterial colonies containing the construct.

Prior to use of the plasmid construct, confirmation of the presence of the cj1411c gene was required. Performing a double digest with BamHI and NdeI restriction enzymes produced a band of ~1.3 kbp in size, which corresponds to the predicted size of the cj1411c gene (1362 bp), confirming its presence within the pET15b plasmid (Figure 4.1). A sample of non-digested plasmid was sent for sequencing (Source Bioscience LifeSciences, Nottingham) to verify that the gene sequence was correct and that there were no mutations present. The pET15b/cj1411c construct was then transformed into E. coli XL-1 Blue cells in order to produce a stock of the construct for the duration of the project.
Figure 4.1. Restriction digests of pET15b/cj1411c plasmid construct. Uncut plasmid migrated relatively slowly through the gel (C). A single cut with either BamHI (B) or NdeI (N) linearised the plasmid and it ran with an apparent size ~7.0 kbp. With a double digest using both enzymes B + N, the cj1411c gene was excised from the vector and ran at ~1.3 kbp. The empty pET15b vector ran to ~5.7 kbp. Samples were run on a 0.8 % agarose gel with a 2-Log DNA ladder (M) present to help identify band sizes.
4.3. Expression trials of CYP172A1 in multiple *E. coli* strains

The pET15b/cj1411c plasmid construct was transformed into five different *E. coli* strains: BL21 (DE3), BL21RP (DE3), HMS174 (DE3), Origami (DE3) and Rosetta (DE3). These five strains are the mutated progeny from different ancestral bacteria (*E. coli* K-12 and B834). The Rosetta cell line also contains enhanced rare codons, which aids the translation of some proteins. The selected strains were used to select for the highest expression of CYP172A1. The five strains all include the DE3 lysogen, which contains the T7 RNA polymerase gene required by the plasmid to allow expression to occur from the T7 lac promoter. This gene was under the control of the lacUV5 promoter, which was regulated by the inducer IPTG (Isopropyl β-D-1-thiogalactopyranoside).

Initially, to investigate CYP172A1 expression, the plasmid-containing *E. coli* cells were grown in 50 ml liquid cultures. These transformants were grown at 37 °C until they reached an OD$_{600}$ of ~0.6, at which point half of the cultures were moved to an incubator set to 25 °C. Within these two different growth temperatures, CYP172A1 expression in some of the flasks was induced with IPTG (1 mM), while IPTG was not added to others (to look at baseline expression of CYP172A1 in the absence of an inducer). Samples from all the flasks were removed at $t = 0$, 3, 5 and 16 hours to examine time-dependent expression of CYP172A1. These samples were then analysed by SDS-PAGE and a band at 54.6 kDa (the predicted size for CYP172A1) was seen to increase over time (Figure 4.2).

In *E. coli* HMS174 (DE3) and *E. coli* BL21 (DE3), there were low levels of CYP172A1 expression in the absence of IPTG. However, in both instances the induced cells produced large quantities of the recombinant protein. For all of the samples tested, as the time increased the amount of CYP172A1 produced also increased, which indicated successful induction of the protein (data not shown).

CYP172A1 expression was successfully induced, and the P450 detected in all of the strains tested, but the solubility of the protein was unknown at this stage as only the total protein content had been studied (data not shown). To investigate this further, cell samples were mixed with a BugBuster® solution to lyse the cells. An aliquot was
removed from this suspension (to be used to define the total protein content). The remaining solution was then centrifuged in a bench top centrifuge, and the supernatant (S/N), that contained the soluble fraction of the sample, and the pellet containing the insoluble fraction were separated. These three fractions were analysed by SDS-PAGE. *E. coli* HMS174 (DE3) was the only strain to express soluble CYP172A1 (Figure 4.3). At 37 °C there was no soluble P450 produced from the induced sample after five hours. However, at the 25 °C induction temperature, soluble protein was produced, but in a roughly equal amount to the insoluble material. As only HMS174 (DE3) appeared to produce any significant amount of soluble protein, this was the only cell line that was carried forward for further optimisation of conditions in order to improve the protein yield.

### 4.3.1. Investigating the effects of varying the temperature at which CYP172A1 expression is induced

In 50 ml cultures, *E. coli* HMS174 (DE3) cells were grown at 37 °C to an OD$_{600}$ of 0.6 and then cooled to 18 °C and 23 °C prior to induction with IPTG (1 mM). Samples were removed and P450 production analysed in a time dependent manner at the two lower growth temperatures. There were no observable differences in the amounts of soluble CYP172A1 produced at these two temperatures, but they both showed an improvement compared to cells grown and induced at 37 °C (Figure 4.4). Some expression of CYP172A1 occurred in the absence of inducer at the lower temperature, but higher levels of CYP172A1 were produced on induction with IPTG. This suggests slightly “loose” expression regulation using the pET15b vector. Lower temperatures are frequently found to promote protein solubility in *E. coli* since slower growth and gene expression may be more conducive to correct protein maturation. Higher growth temperatures may increase the amount of misfolded proteins, which can either be degraded or otherwise aggregate to form inclusion bodies [235].
4.3.2. Analysis of the effects of varying IPTG levels and the growth medium used to express CYP172A1

Varying the concentration of IPTG during induction (between 50-500 µM) had no apparent effect on the levels of production of CYP172A1 after 16 hours post induction, nor did it change the apparent ratios of soluble to insoluble protein to any significant extent, whereby the majority of the induced protein was in the soluble fraction (data not shown). The concentration of IPTG was therefore maintained at 1 mM in subsequent studies. Three different growth medium types, Luria Bertani (LB), Terrific Broth (TB) and Yeast Tryptone (2xYT), were used to see if any improvements could be made to the final yield. Analysis of SDS-PAGE gels measuring the total protein vs. soluble protein after 16 hours indicated that 2xYT was the medium that produced levels of soluble CYP172A1 greater than LB and TB (data not shown) and was subsequently used for further expression of the P450.
Figure 4.2. Time dependent production of CYP172A1. A time course monitoring protein expression levels in E. coli HMS174 (DE3) at t = 0 (lane 1), t = 2 (lanes 2 and 5), t=5 (lanes 3 and 6) and t=16 (lanes 4 and 7) hours at 37 °C induction temperature: and at t =2 (lanes 8 and 11) t=5 (lanes 9 and 12) and t=16 (lanes 10 and 13) hours at 25 °C induction temperature. Red numbers indicate samples from cells to which IPTG (1 mM) was added. Blue numbers indicate samples from cells to which no IPTG was added. Samples were run on a 12 % SDS-PAGE gel and band sizes were compared to a protein marker (M). A band appeared over time at the expected size for CYP172A1 (54.6 kDa).

Figure 4.3. Solubility of CYP172A1 produced in E. coli HMS174 (DE3). The gel (lanes 1-3) shows the total amount of cell protein produced in 5 hours (lane 1), along with the proteins found in the soluble fraction (lane 2) and in the insoluble fraction (lane 3) from cells grown at 37°C. Lanes 4-9 show the total amount of protein produced in 16 hours (lanes 4 and 7), along with the proteins in the soluble fraction (lanes 5 and 8) and in the insoluble fraction (lanes 6 and 9) from cells grown at 25°C. Red numbers indicate cells in which CYP172A1 was induced with IPTG, while blue numbers indicate uninduced cells. Samples were run on a 12 % SDS-PAGE gel with a protein marker (M). A strong band at the correct size for CYP172A1 was observed (~54.6 kDa).
Figure 4.4. Effect of induction growth temperature on the solubility of CYP172A1 produced in E. coli HMS174 (DE3). The gel indicates the total amount of cell protein produced in 5 hours (lane 1), the proteins present in the soluble fraction (lane 2), and in the insoluble fraction (lane 3) from cells grown at 23 °C. In lanes 4-9, the total amount of protein produced in 16 hours (lanes 4 and 7), and in the soluble fraction (lanes 5 and 8) and in the insoluble fraction (lanes 6 and 9) is shown from cells induced at 23 °C. Red numbers indicate cells in which CYP172A1 was induced with 1 mM IPTG, while blue numbers indicate uninduced cells. Samples were analysed on a 12 % SDS-PAGE gel with band sizes estimated using a protein marker (M).
4.4. Large scale expression and purification of CYP172A1

Following the preliminary work described above, *E. coli* HMS174 (DE3) transformed with the pET15b/cj1411c plasmid construct was used to produce CYP172A1 for all subsequent experiments. Samples were taken throughout the expression and purification stages, and analysed by SDS-PAGE to examine expression levels, the solubility and purity of CYP172A1 (Figure 4.5). Approximately 12 l of culture was induced and grown at 25 °C for 16 hours. Cells were centrifuged, at 5,000 g, and the pellet resuspended in protein binding buffer (100 mM KP, with 250 mM KCl and 10% (v/v) glycerol at pH 7.5). The cells were lysed using a continuous flow French Press at 950 lb/in² with the material kept cold with ice. Lysed cell debris was removed by ultracentrifugation (40,000 g for 35 minutes at 4 °C) and the supernatants pooled. A two step purification method was applied to purify CYP172A1. The protein was first loaded onto a Ni-NTA column and, by increasing the concentration of imidazole in the protein binding buffer, contaminating proteins were progressively removed. Some elution of CYP172A1 began at 150 mM imidazole, but this fraction still contained other unwanted proteins, and so only the P450 that eluted in the 250 mM imidazole fraction was kept for further purification. Following dialysis, to remove the imidazole bound to the P450, CYP172A1 was concentrated in a Vivaspin 20 (Mₜ cut off at 30 kDa). The concentrated CYP172A1- containing sample (typically 2-5 ml) was then loaded onto a Sephadex S200 size exclusion column preequilibrated with either 100 mM Tris buffer with 250 mM KCl at pH 7.5, or 50 mM KP_i buffer with 200 mM KCl at pH 7.5 (Section 2.29). This purification step produced CYP172A1 which, when analysed spectrophotometrically had a 418/280 nm (R₂) ratio of at least 1.3. Following electrophoresis of samples isolated following gel filtration, a dominant band was observed at the expected Mₜ (~54.6 kDa) for CYP172A1 (Figure 4.6), and mass spectrometry analysis using Mascot confirmed the identity of this protein band as CYP172A1, with a mass of 52.6 kDa. The masses differ by the size of the addition His₆ tag on the recombinant protein. On the gel two minor bands below CYP172A1 are observed, which possibly reflect some minor degradation of the P450. This degradation may compromise the homogeneity of the protein sample, which might affect the likelihood of crystal formation.
The CYP172A1 elution profile observed on the S200 column was very broad and a large proportion of the P450 eluted with the void volume (Appendix, Figure S4), suggesting a very high molecular weight for some of the P450. Bands of higher mass ~100 kDa and 150 KDa (blue arrows Figure 4.6) than the expected 54.6 kDa were also often observed on SDS-PAGE gels, which might represent dimer and trimer, respectively. These high molecular weight bands were analysed by mass spectrometry and were also identified as CYP172A1 based upon their peptide signature in Mascot (Appendix, Figure S5). These CYP172A1 protein aggregates were apparently stabilised by interactions which were quite resistant to denaturation by SDS, DTT and boiling. Despite the formation of these high M\text{w} structures, the solubility of the P450 was apparently unaffected. It has been reported that some proteins can form aggregates but remain soluble in solution [236]. Investigations into this aggregation phenomenon are described in Section 4.13.
Figure 4.5. Production and Ni-NTA purification of CYP172A1 from *E. coli* HMS174 (DE3). Cells were grown to an OD$_{600}$ of ~0.6 (37 °C) and then gene expression was induced with 1 mM IPTG at 25 °C. At t=0 (lane 1) no CYP172A1 was detected in the HMS174 (DE3) transformants. After 16 hours a band at the expected M$_w$ is observed (lane 2). Lysed cells were centrifuged leaving soluble (lane 3) and insoluble material (lane 4). The soluble material was passed through Ni-NTA resin and non-specific protein did not bind (lane 5). The resin was washed with buffer (lane 6) and 50 mM imidazole (lane 7) to remove unwanted proteins. CYP172A1 was eluted with 250 mM imidazole (lane 8) with an expected M$_w$ of ~54.6 kDa. A marker (M) indicates the masses of the protein bands. Samples were run on a 10 % SDS-PAGE gel.

Figure 4.6. Size exclusion purification of CYP172A1. CYP172A1 fractions eluting from a size exclusion Sephadex S200 column (lanes 1-8) show much increased purity compared to the protein sample used for column loading (pre). Sample fractions were run on a 10 % w/v SDS-PAGE gel with a marker (M) to determine band sizes, CYP172A1 is shown (red arrow) with an expected M$_w$ of 54.6 kDa. Higher molecular weight bands that were also identified as CYP172A1, by mass spectrometry, are marked (blue arrows).
4.5. Initial spectroscopic characterisation of CYP172A1

Having produced and purified a red protein, it was then vital to identify it as a heme-containing P450 protein. This was achieved by examining the protein’s spectroscopic properties. CYP172A1, in its resting (oxidized) state, produced a typical UV-visible spectrum for a ferric, inhibitor/substrate free P450 with the main Soret peak at 418 nm (black spectrum, Figure 4.7). The enzyme was in a predominantly low spin (LS) state, due to the apparent absence of a peak ~390 nm, which would indicate a proportion of the enzyme being in a high spin (HS) ferric state [237]. The lower intensity α, β and δ bands, typical of a heme protein, were all clearly visible at 569, 535 and 355 nm, respectively. Upon the addition of a few grains of solid sodium dithionite reductant, no obvious changes were observed to the spectral properties of the enzyme, except that the δ peak was lost under the absorbance from the dithionite itself (green spectrum). This indicated that, under the aerobic conditions used, CYP172A1 heme iron was not significantly reduced on the addition of dithionite. The subsequent addition of CO to CYP172A1 produced a shift in the Soret peak to 448 nm as the CO bound to the reduced (ferrous) heme iron to create the P450 Fe$^{2+}$-CO complex (red spectrum) [238]. This suggests that, although the CYP172A1 remained predominantly ferric upon the addition of dithionite, a small proportion of the heme is reduced to the ferrous state, and the reduced heme in this state rapidly binds to CO. Over time, the equilibrium between the ferric and ferrous form is pulled toward the latter species, which binds CO and leads to the near-total formation of the stable Fe$^{2+}$-CO complex. In the absence of the CO it is likely that the ferrous heme iron instead reacts efficiently with oxygen, and is converted back to the ferric heme iron, creating superoxide in the process. The Fe$^{2+}$-CO complex with Soret peak at 448 nm is consistent with the iron ligand trans to the CO being cysteine thiolate from Cys399. A small shoulder is visible at ~420 nm in the Fe$^{2+}$-CO bound form, and this likely occurs since a small proportion of the cysteine thiolate becomes protonated to create a thiol-ligated CO form, which is known as the P420 complex [186]. The α and β bands spectrally merge together in the ferrous state, and a new broad peak is formed at ~545 nm in the visible region. The binding of NO to ferric CYP172A1 forms a Fe$^{3+}$-NO complex and the Soret peak blue shifts to 432 nm, a type of absorbance shift that is reported for other P450s [239] (blue dash, Figure 4.7).
was also a shift in the position and an increase in intensity of the α and β peak bands, and these features become more resolved in the NO complex.

4.6. Calculation of the molar extinction coefficient for CYP172A1 using the pyridine hemochromagen determination method

Through studies done on multiple P450s it has become apparent that use of a single molar extinction coefficient, as described by Omura and Sato, of 91 mM$^{-1}$ cm$^{-1}$ for the absorbance difference ($\Delta A_{450} - \Delta A_{490}$) between Fe$^{2+}$-CO and Fe$^{2+}$ forms of these enzymes is not accurate in all cases. This is due to innate differences in the heme environments, absorbance maxima, P450/P420 equilibria etc [186]. The quantification of the heme extinction coefficient was thus conducted using the pyridine hemochromagen method, as set out by Berry and Trumpower [185]. The change in the absorbance between the oxidised (red spectrum, Figure 4.8) and the dithionite-reduced heme-pyridine complex (green spectrum) in the Q-band region was used. The extinction coefficient for the heme-pyridine complex has been previously estimated by Berry and Trumpower and, by rearranging the Beer-Lambert rule, the molar extinction coefficient was determined for the oxidized/ferric CYP172A1. From three independent measurements, the extinction coefficient for CYP172A1 was determined to be $\varepsilon_{418} = 105 \pm 5$ mM$^{-1}$ cm$^{-1}$ for the resting ferric form of the P450. This value was used whenever a concentration based on the Soret peak was measured. This value is lower when compared to coefficients reported for certain other P450s determined using this method, including CYP121 (134 mM$^{-1}$ cm$^{-1}$) and CYP142 (140 mM$^{-1}$ cm$^{-1}$) both from Mycobacterium tuberculosis [240, 241], but is similar to the extinction coefficient reported for the P450 BM3 heme domain (95 mM$^{-1}$ cm$^{-1}$) and for CYP144 (M. tuberculosis) (100 mM$^{-1}$ cm$^{-1}$) [242, 243].
Figure 4.7. Spectroscopic characterisation of ferric, ferrous, ferrous/CO-bound and ferric/NO-bound CYP172A1. The UV-Visible absorption spectra of CYP172A1 (~3 μM) in the resting ferric LS Cys-H2O ligated (black); the dithionite treated, but still predominantly ferric (green); the ferrous CO-bound (red); and the ferric NO-bound (blue dash) states are shown. Soret absorption maxima are located at 418, 418, 448 and 432 nm, respectively. Distinct α and β bands are visible in the ferric spectrum at 569 and 537 nm. These spectrally merge into a single broad peak at 552 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 432 nm and more prominent α and β peaks emerge, when compared to the oxidized enzyme, at 573 and 543 nm, respectively. All four species were characterized under aerobic conditions with CYP172A1 in protein binding buffer.
Figure 4.8. Pyridine hemochromagen complex formation of CYP172A1. UV-Visible spectra are shown for the native (ferric) CYP172A1 (2.43 μM, black spectrum), following formation of the pyridine hemichromagen complex (red spectrum), and the reduction of the complex by addition of sodium dithionite (green spectrum). Insert, shows a magnified view of the spectral changes between occurring 500 and 600 nm. The pyridine hemochromagen complex has a maximum absorbance at 556 nm.
4.7. Determination of the heme coordination state in CYP172A1 by Electron Paramagnetic Resonance (EPR) spectroscopy

EPR is a technique that is able to detect the environment of unpaired electrons in samples, and to quantify and identify various radicals present. This makes it a very good tool for looking at P450s as they contain one or more unpaired electrons in the heme iron in the oxidised (ferric) form. EPR can thus be used to study the oxidation state of the heme iron, including if it is in a LS or HS ferric state. The EPR spectra for these two states are distinct and have diagnostic g factors (or g values) [244]. The EPR signal of native CYP172A1 indicated that the protein is in a single LS state with g values of 2.41 (g_z), 2.24 (g_y) and 1.91 (g_x) (Figure 4.9). These g values are consistent with the heme being proximally coordinated by a cys-thiolate species, and with a distal water ligand in the 6th position. These g values are very similar to those reported for other P450s e.g. P450cam (2.45, 2.26, 1.91), CYP121 (2.47, 2.25, 1.90) and BM3 (2.42, 2.26 and 1.92) [141, 245]. The homogenous LS state of the CYP172A1 heme iron may reflect a lack of flexibility in the axial ligands and a single major conformational state of the heme iron. Binding of ligands altered the observed g values (see Section 5.5).

4.8. Analysis of the tertiary structure of CYP172A1 using fluorescence spectroscopy

The P450 enzymes are often not particularly stable proteins, and have a tendency to form the P420 (cysteine thiol-coordinated) form when exposed to denaturants, high temperatures or pressures [246]. To investigate the stability of the tertiary structure of CYP172A1 to a chaotrope, increasing concentrations (0-4 M) of guanidine hydrochloride (GdnHCl) protein denaturant were added to CYP172A1, and tryptophan fluorescence monitored on incubation across a range of GdnHCl concentrations. A stable protein usually buries hydrophobic aromatic side chains into the core of the protein away from the solvent, an energetically favourable process [247]. As a consequence, there is often only low fluorescence emission from excited tryptophans, as well as other aromatic amino acids, as energy emitted is lost to nearby residues through vibrations and other processes. However, as the protein’s tertiary structure is disrupted
the aromatic residues become more solvent accessible, and so their fluorescence emission intensity upon excitation often increases as the protein unfolds.

As increasing concentrations of GdnHCl were added to CYP172A1, the tertiary structure was seen to be disrupted from fluorescence changes, and this also resulted in the release of the bound heme and the loss in absorbance intensity of the thiolate-coordinated heme Soret peak (Figure 4.10). A new Soret peak appeared at 370 nm in a manner similar to that occurring on denaturation of P450 BM3[248]. The dissociation of the heme cofactor and the formation of 370 nm peak are associated with heme loss from the protein scaffold, and the breakage of its bond to cysteine thiolate [143]. At increasing concentrations of GdnHCl the tryptophan emission increased, with the fluorescence emission peak shifting from 325 nm to 352 nm. This is interpreted as reporting directly on the loss of tertiary structure as tryptophans (and other aromatic amino acid side chains) become more accessible to the solvent. The change in fluorescence emission from CYP172A1 tryptophan residues at 352 nm was then plotted against the corresponding concentration of GdnHCl, and the data were fitted using the Hill equation to give a midpoint of 1.30 ± 0.1 M for tertiary structure loss. Compared to CYP175A1 (2.6 M) from *Thermus thermophilus* HB27, the CYP172A1 value was quite low. However, CYP175A1 originates from a thermostable organism and has structural adaptations that enable it to tolerate high temperatures (and chaotropes) – primarily through salt bridge networks in the protein and the small of this P450 [249, 250]. Ligand free *B. subtilis* P450 BioI (CYP107H) has a similar structural stability in the presence of GdnHCl (1.5 ± 0.2 M), suggesting that this mesophilic bacterium P450’s tertiary structure stability is comparable to that for CYP172A1 [143].
Figure 4.9. EPR spectrum of native CYP172A1. The X-band EPR spectrum shows the oxidised (ferric) CYP172A1 to exist in a single low spin state with g-values at 2.41, 2.24 and 1.91. This spectrum is consistent with that for a LS state P450 heme proximally coordinated by a cysteine thiolate ligand.
Figure 4.10. Analysis of the stability of heme binding and tertiary structure in CYP172A1 on exposure to increasing concentrations of GdnHCl. (A) Selected UV-Visible absorption spectra of CYP172A1 (1.5 µM) with increasing concentrations of GdnHCl (0-4 M). The resting Soret peak (black spectrum, 419 nm) decreased in intensity with increasing concentrations of GdnHCl (0.75 M – red dotted spectrum, 1.25 M – green dotted spectrum, 1.65M - blue dotted spectrum, 2.25 M – cyan dotted spectrum, 4.0 M – pink spectrum) and shifted to 370 nm, indicative of loss of heme ligation and heme dissociation from the protein. Insert, change in tryptophan fluorescence as the [GdnHCl] increases. (B) A plot showing the percentage increase in tryptophan emission at 352 nm (relative to the fluorescence value for the resting form of the protein) against the concentration of GdnHCl. The concentration where 50 % of the protein was unfolded was calculated, using the Hill function, to be 1.30 ± 0.1 M.
4.9. Analysis of the secondary structures of CYP172A1

Circular dichroism (CD) is a technique that looks at the differences in the absorbance of left and right rotations of polarised light. This phenomenon can be used to identify and quantify stereoisomers, such as those for L- and D-amino acids. Far UV (FUV) CD spectroscopy can be used to examine the secondary structure in a protein and to provide estimations of the proportions of the protein in α helical, β sheet and other secondary structural conformations [251]. These secondary structures are highly ordered and exist in a stable environment. Commonly, an α helix will give diagnostic signals (spectral minima) at ~208 and ~222 nm, whereas the β sheet gives a CD minimum of weaker intensity at ~217 nm [252]. The strength of the signals will decrease if the protein’s integrity is compromised, e.g. if a protein denaturant, such as GdnHCl is added. The more unfolded the protein becomes, the more the FUV CD signal is diminished, and this is ultimately lost completely when all the secondary structure is disrupted.

The FUV CD spectrum for CYP172A1 indicated that the P450 was predominantly alpha helical (black spectrum, Figure 4.11), which would be expected as the solved crystal structures of other P450s are all mainly alpha helical [117]. The spectrum contained a small positive CD peak at 198 nm and two negative CD minima at 210 and 222 nm. Data points from the spectrum were inputted into the k2d program (http://kal-el.ugr.es/k2d/spectra.html) and the α helical content was calculated to be 60 %, the β sheet content 7 % and the proportion of random coil 33 %. Compared to other P450s the value for the alpha helix content is similar, with the majority of P450s typically having >50 % alpha helix [253]. However, highly accurate predictions of secondary structural content based on CD are not often possible, and rely on high quality far UV CD spectra that extend to wavelengths below 190 nm.

With increasing concentrations of GdnHCl, the FUV CD spectrum for CYP172A1 was diminished. With the addition of 2.25 M GdnHCl the P450 appeared to be almost completely denatured, since the spectrum was quite extensively lost. However, at this stage the spectral data became affected by an increase in turbidity (caused by protein aggregation), making it difficult to collect useful CD data at higher concentrations of GdnHCl. The k2d program predicts only 2 % alpha helix, 51 % beta sheet and 47 %
random coil at 2.25 M GdnHCl. However, this result likely overestimates that proportion of beta sheet considerably.

4.10. Measuring CYP172A1 thermostability using Differential Scanning Calorimetry (DSC)
Differential Scanning Calorimetry (DSC) is a method that looks at thermal changes that occur within a protein upon its heating, and is used to measure the changes in enthalpy and entropy upon protein unfolding. Despite multiple attempts to obtain DSC data, CYP172A1 was found to aggregate and precipitate prior to reaching a temperature when any significant amount of protein unfolding occurred, and thus no useful results were available to analyse.

4.11. Analysis of the molecular weight and polydispersity of CYP172A1 by MALLS
Multiangle laser light scattering (MALLS) is a technique that works by measuring the amount of light scattered as a sample passes through a detector. Light scatter is dependent on the concentration of the solute and its molecular mass, and therefore it is a good method for measuring the $M_w$ of a sample. Information about the homogeneity is also available, which can help to determine the suitability of a sample for protein crystallography. This is measured by a polydispersity index, which indicates whether aggregates or multiple forms (monomer, dimer etc) of the sample are present. MALLS was carried out on purified CYP172A1 to determine the oligomeric state of the protein. Prior to MALLS analysis the sample was passed down a size exclusion column, which fed directly into the MALLS instrument. Under non reducing conditions a single peak was observed, which eluted from the column ~13 ml (Figure 4.12). Despite the presence of a single peak the polydispersity value did not equal one, which indicated CYP172A1 did not exist as a single, uniform species. Across the peak a range of $M_w$’s were observed. At the start of the peak CYP172A1 was predicted as an approximate 10-mer (~500 kDa) and at the end of the peak an approximate tetramer (~200 kDa). The result indicated that there were multiple oligomeric forms of the P450 present. This result was
unexpected as the majority of bacterial P450s exist in a monomeric form, with a polydispersity index equal to or close to one [254].

Repeated efforts using various buffer conditions were made. However, the MALLS data consistently indicated a polydispersity value above one and across the peak multiple molecular weights were observed, indicating that CYP172A1 likely self-aggregates. The results indicated that, in the aggregated state(s), protein crystallogenesis might prove difficult to achieve, as monodispersity is a key factor for successful crystallography [255].
Figure 4.11. Effect of GdnHCl on the FUV CD spectra of CYP172A1. Native CYP172A1 (black spectrum) shows a CD maximum at 198 nm and two minima at 210 and 222 nm. Additions of a range of [GdnHCl] 0.25 M (green dotted spectrum), 0.5 M (red dotted spectrum), 0.75 M (blue dotted spectrum), 1.25 M (cyan dotted spectrum) and 2.25 M (magneta spectrum) progressively decreased the FUV CD intensity of each spectrum, indicative of increasing loss of secondary structure.

Figure 4.12. MALLS analysis of native CYP172A1. The figure shows a plot of the refractive index and molecular mass versus the elution volume for a MALLS experiment with CYP172A1. A single broad peak was observed from the light scattering. Across the peak a range of molecular masses was observed (4.7 x10^5 - 2 x10^5 Da). The experiment was run in a 10 mM Tris buffer, with 100 mM KCl at pH 7.
4.12. Use of the Thermofluor probe to identify stabilising buffers for CYP172A1

Further investigations into the nature of the oligomeric forms of CYP172A1 started with the identification of a buffer that would decrease the aggregation of CYP172A1, and which thus might stabilise the protein and increase its melting temperature (T_m). The thermofluor assay has been used for other proteins as a way to complement MALLS analysis in determining the T_m and any conditions that improve this parameter, and as a tool for identifying appropriate buffer conditions before initiating crystallisation trials [256, 257]. The experiment allows for large scale screening of multiple buffer conditions, such as those present in commercially available crystallisation screens. The use of the scanning function on a real time PCR machine (Bio-Rad) allows for fluorescence detection from the SYPRO dye (Ex 462 nm and Em 569 nm) that is added to each of the sample wells (Section 2.38). The fluorescence of the dye increases when it is exposed to hydrophobic environments [258]. This means that, as the P450 is unfolded and the buried residues are exposed, there is an increase in SYPRO fluorescence that associated with dye binding to exposed hydrophobic regions. The midpoint of the unfolding transition can then be derived from the fluorescence melting curve to calculate the T_m of the protein.

CYP172A1 was first screened using the JBS solubility screen kit, which covers a pH range of 3-10 with a range of different buffers, including common buffers such as KP_i, HEPES and Tris, as well as less commonly used buffers, such as bicine, CHAPS and PIPPS (all at 50 mM working concentrations). A wide range of CYP172A1 T_m values were observed with this initial screen (Figure 4.13). Buffers at low pH (below five) produced T_m values at temperatures below 40°C and were non-stabilising. The majority of the buffers produced T_m peaks between 45-55 °C (Appendix, Table S2), but there were two buffers that appeared to maximally stabilise the protein: these being, Na/K phosphate (pH 7.0) and imidazole (pH 8.0). These buffers produced CYP172A1 T_m’s of 56.4 and 57.0 °C, respectively. The increase in the observed T_m for the imidazole buffer likely occurred as P450s can be stabilised by the presence of ligands, and imidazole binds in a type II manner to P450s. For further stability testing the Na/K phosphate buffer was then used in combination with the JBS additive screen.
Chapter 4. Expression, purification and biophysical characterisation of CYP172A1

Figure 4.13. Thermofluor assay for CYP172A1 using the JBS buffer screen. The effect of several buffers (covering a range of pH’s) on the T_m of CYP172A1 are displayed. The graph shows the first derivative of the change in fluorescence emission over the temperature range. The T_m was calculated from the observed peak. Conditions where the peak appeared at higher temperatures had a more stabilising effect on the P450. Very low pH’s (<3.2) failed to produce a peak, which suggests that the protein unfolded in advance of initiating the temperature gradient.
CYP172A1, prepared in the 50 mM K/Na phosphate buffer (pH 7.0), was used along with the JBS additives screen. The screen consists of 14 additives including reducing agents (DTT), salts (NaCl) and detergents (Dodecyl maltoside). CYP172A1 (6 mg/ml) was added to each buffer and the T_m’s determined in each case (Figure 4.14 and Appendix, Table S3). CHAPS and TCEP had negative effects of the stability of the protein and lowered the T_m compared to phosphate buffer control. The maltosides failed to produce a melting temperature, and so were also not suitable additives. Sodium chloride increased the T_m by 0.6 °C to 57.0 °C. The additive which was the most stabilising was glycerol, which produced T_m’s of 57.2 °C (20 % glycerol) and 57.7 °C (40 % glycerol).

Salts and glycerol stabilisation effects on proteins have been previously reported [259-261]. For crystallography, a high concentration of glycerol can have a negative impact on crystal formation. The glycerol may also stabilise the protein in solution so that it does not precipitate out to form a crystal under certain conditions that might otherwise promote crystallogenesis. Salts can also be a problem as they have a tendency to form their own crystals. However, when crystallising co-factor-containing proteins, such as P450s, their coloured nature means that they can be distinguished from colourless salt crystals, and so salt crystallisation is less of an issue with respect to confusing salt crystals with protein crystals.
Figure 4.14. Thermofluor assay for CYP172A1 in Na/K phosphate buffer with JBS additives. Thermofluor unfolding data were examined for the effects of additives to Na/K phosphate buffer (pH 7.0) (buffer 14) on the $T_m$ of CYP172A1. The graph shows the first derivative of the change in fluorescence emission over the temperature range. The $T_m$ was calculated from the observed peak. Conditions where the peak appears more to the right (higher temperature) have a better stabilising effect on the protein. Some additive/CYP172A1 mixtures failed to produce a peak, which suggests that the protein was unfolded already in these conditions prior to initiation of the temperature gradient.
4.12.1. Analysis of the effect of ligand binding on the thermostability of CYP172A1

From the first thermofluor stability assay, it was noted that a 50 mM imidazole buffer produced a higher T_m compared to the other buffers. This is most likely due to the ability of imidazole to act as a 6th ligand to the heme iron of CYP172A1. To explore this further, 1-Phenylimidazole (1-PIM) was chosen in order to study its effects on the thermostability of CYP172A1, as it was shown to bind tightly to CYP172A1 (Section 5.2.2.1) and is readily soluble. The Na/K phosphate buffer was chosen and the 1-PIM bound protein was screened using the additives kit.

The addition of the 1-PIM had a beneficial effect on the stability of the protein (Figure 4.15). Comparing the T_m’s of the unbound protein (dotted line) to the 1-PIM bound protein (solid black line) indicated that there was an increase in the T_m from 56.4 °C to 59.8 °C. The additives screen indicated a similar pattern of results with 1-PIM-bound CYP172A1 compared to the unbound protein. CHAPS, glucopyranosides and the maltosides had a negative impact on the T_m and the glycerol additives displayed a slight improvement compared to the 1-PIM bound CYP172A1 in the standard buffer. Another feature of the ligand bound protein was the increase in the sharpness of the peaks used to calculate the T_m’s, compared to the previous experiments in the absence of 1-PIM, where the peaks were much broader.

The results generated in the thermofluor assay indicated that the most stabilising conditions for CYP172A1 were a 50 mM phosphate buffer (pH 7.0) supplemented with glycerol, and that adding salts had a slight positive effect on the protein’s stability. The addition of a ligand (1-PIM) greatly improved the stability, producing a 3.4 °C increase in the melting temperature. As a consequence of these data attempts at crystallisation of CYP172A1 were conducted in the phosphate buffer, both with and without the 1-PIM ligand at 10x the K_d concentration (determined in Section 5.2.2.1).
Figure 4.15. Thermofluor analysis for 1-PIM-bound CYP172A1 in Na/K phosphate buffer with JBS additives. The effects of the listed additives to Na/K phosphate buffer on the $T_m$ of 1-PIM bound CYP172A1 are shown. The graph shows the first derivative of the change in fluorescence emission over the temperature range. The $T_m$ was calculated from the observed peak. Conditions where the peak moves to the right (higher temperature), with respect to that for the additive-free CYP172A1 have a better stabilising effect on the protein.
4.13. Investigations into the causal effects of CYP172A1 self aggregation
As there is a tendency for CYP172A1 to self aggregate, various experiments were conducted to investigate this property, with the hope of resolving the issue to enable the generation of a monomeric species, as occurs with most of the other bacterial P450s, and to allow for successful crystallography screening [254]. Aggregation of proteins can be mediated by a number of different interactions, and typically involve hydrophobic, Van der Waals, disulphide bridges and ionic interactions. Aggregation is often caused by exposed hydrophobic regions on proteins interacting with each other [262]. This initial contact can then be strengthened by other favourable interactions (e.g. involving ionic interactions or disulphide bridges).

4.13. Effects of reducing agents on CYP172A1 aggregation
In an attempt to decrease the aggregation and oligomerisation of CYP172A1, the influence of various protein structural “disrupters” was investigated. In particular, the reducing agents DTT, BME and low concentrations of the protein denaturant GdnHCl were added to solutions of CYP172A1. The most effective of these chemicals was BME (at 100 mM), as confirmed by size exclusion chromatography (Appendix, Figure S6) and SDS-PAGE (Figure 4.16). The early fractions eluting from the column of BME-treated CYP172A1 contained the high M_w bands, which were previously confirmed as CYP172A1 aggregates by mass spectrometry (Section 4.4). However, the fractions collected from the large peak which eluted from the size exclusion column after ~50 minutes (15 ml) revealed decreasing amounts of these aggregated species. The 418:280 absorption ratios from the peak samples were ~1.3, very similar to the ratios previously found (Section 4.4), and so it was probable that this new peak contained a greater proportion of a less aggregated protein, rather than a more purified CYP172A1 sample.

The C1 and C2 fractions from the size exclusion separation were combined and sent for SEC-MALLS analysis (Figure 4.16). The column was equilibrated with the same buffer as the protein, except that the concentration of BME was decreased to 1 mM, to prevent interference with the MALLS detector. The protein sample produced a peak that was ~2 ml wide. The main peak was centred at 15 ml (identical to that for the elution from the
S200 column) with a $M_w \sim 47.8$ kDa, which was a little lower than predicted ($M_w$ of CYP172A1 is 54.6 kDa). It is unclear why the two values are not more similar, as the accuracy of MALLS is often ± 1% of the actual $M_w$ [263]. Adjacent to the main peak, two smaller peaks, one of a slightly larger mass (shorter elution time) and one of a smaller mass (longer elution time) were present. These may represent minor contaminants, or possibly a small proportion of proteolysed CYP172A1 (for the later peak). Because of the apparent decrease in the levels of aggregation in the CYP172A1 protein samples with 100 mM BME present, BME was subsequently used in trials for the preparation of crystals (Section 4.16).
Figure 4.16. SDS-PAGE analysis of S200 fractions and MALLS analysis of CYP172A1 purified in the presence of 100 mM BME. A) 12% SDS-PAGE gel of fractions obtained from a size exclusion gel filtration column. A major band is observed at the correct Mₗ for CYP172A1 (~54 kDa) compared to the protein marker (M). Fractions C1 and C2 appear to have lost the higher molecular weight species that may be associated with aggregated forms of CYP172A1 (blue arrows). B) MALLS analysis of the combined C1 and C2 fractions. A peak was observed from the light scattering eluting between 14-16 ml. Across the peak a range of molecular masses (4.7 x10⁴-2 x10⁵ Da) are predicted. The experiment was run in a buffer of 100 mM KP₆, 100 mM KCl and 1 mM BME at pH 7.5.
4.14. Cleavage of the N-terminal His\textsubscript{6} purification tag from CYP172A1

Another possible explanation for the aggregation of CYP172A1 is that it is driven by the His\textsubscript{6} purification tag on the P450, which is used during the initial purification stages. The pET15b vector contains a thrombin cleavage site which was translated into the recombinant CYP172A1 between the tag and the start of P450 coding sequence. In accordance with the manufacturer’s protocol, small scale trials were undertaken in order to determine the optimum conditions for thrombin mediated cleavage of the purification tag. It was determined that for every 10 µg of protein, 0.08 U/ml of thrombin was required for maximum tag removal. This was scaled up for a 300 µg protein sample. CYP172A1 was mixed with the thrombin for four hours at 4 °C and then mixed with a Ni-NTA resin slurry for two hours. The slurry was then poured into a plastic column and the supernatant eluted and kept (SN1). The column was then washed with protein buffer solution and the eluate (SN2) collected.

These two fractions were then analysed by SDS-PAGE, along with non cleaved CYP172A1 (Figure 4.17). The first two lanes indicate a slight increase in CYP172A1 mobility compared to the control, indicating successful cleavage of the purification tag. The protein bands were transferred onto a PVDF membrane, and the membrane probed with an anti-His tag antibody (Figure 4.17). On the western blot, a strong band was observed in the control and a very faint band is seen in SN2, which suggests that a small amount of His-tagged protein was washed off the resin. Visual detection of the anti-His antibody was not observed in SN1, indicating that complete cleavage of the purification tag had occurred. The anti-His tag antibody detected bands with higher M\textsubscript{w}’s in the control sample, but not in the cleaved samples (blue arrow). These bands are also marked in the SDS-PAGE and are also observed in the two cleaved protein samples. This implied that CYP172A1 can form multimers/oligomers regardless of the presence of the His-tag, as bands of higher M\textsubscript{w} were observed in both the cleaved and uncleaved samples of CYP172A1 on the SDS-PAGE gel.
Figure 4.17. Thrombin cleavage of the His\textsubscript{6} purification tag from CYP172A1. A) 12 % SDS-PAGE showing unbound (SN1) and wash through (SN2) from a Ni-NTA column following thrombin incubation, and an undigested control (Uncut). Lanes SN1 and SN2 show a protein with a decreased molecular mass compared to the undigested CYP172A1. B) Western blot using anti-His primary antibody. A dominant band is observed in the “Uncut” lane along with two minor species of higher M\textsubscript{w}, bands for which are absent in lanes 1 and 2 in the western blot, but visible in the SDS-PAGE. The size of the 55.4 kDa protein marker is indicated.
4.15. Mutational truncation of full length CYP172A1

The online program XtalPred was used to predict the likely success of obtaining CYP172A1 crystals, and to look at the effects that N and C terminal truncations might have on improving the likelihood of protein crystallisation. XtalPred prediction software was used on the full length protein (Figure 4.18). The software used multiple variables (protein size, pI, secondary structure etc) to predict the probability of the protein successfully crystallising, based upon databases of protein sequences and their propensity to result in crystal structures. Based upon these variables, XtalPred gave a probability score of 4/5, where a score of one equated to the most desirable and a score of five the least desirable. The isoelectric point (8.96) and the length of CYP172A1 appeared to be the features which contributed to the poor score. A disordered region was also predicted between residues 220 and 238 (a central section of the amino acid sequence). Several truncations at both the C terminus and the N terminus were inputted to see if they could improve the theoretical probability of successful crystallisation.

From trialling several different truncations with XtalPred, three single mutations were chosen, K38H, E109H and S449* (Figure 4.19). The insertion of the histidine residue introduced a NdeI restriction site into the gene, which was then subsequently used to truncate CYP172A1. Mutation S449* introduced a stop codon after the serine residue. Each mutation improved the probability of protein crystallisation to a score of 3/5 (Appendix Figures S7-S9).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Position</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP172A1</td>
<td></td>
<td>453</td>
</tr>
<tr>
<td>S449*</td>
<td>449</td>
<td>449</td>
</tr>
<tr>
<td>K38H</td>
<td>38</td>
<td>453</td>
</tr>
<tr>
<td>E109H</td>
<td>109</td>
<td>453</td>
</tr>
</tbody>
</table>

Figure 4.19. Diagrammatic representation of CYP172A1 and truncated mutant proteins.
The full length CYP172A1 contains 453 amino acids. The locations of each mutation are indicated. Size is not to scale. *= stop codon.
Chapter 4. Expression, purification and biophysical characterisation of CYP172A1

Figure 4.18. Key features used by XtalPred to determine likelihood of CYP172A1 crystallisation. The software calculated the length (453 a.a.), pI (8.96), instability index (44.02), coiled regions (33%) and disordered regions (18 a.a. long) and compared them to all known proteins that have and have not been crystallised. These various factors were combined to give a score of 4/5 for probability of crystal formation. Grey bars = unsuccessfully crystallised proteins, blue bars = successfully crystallised proteins.
Primers to introduce NdeI or stop codon sites in the recombinant *cj1411c* gene were designed and synthesised by Generon. Primer pairs 1/2, 3/4 and 5/6 (Appendix, Table S1) were used to introduce nucleotide changes for the E109H, K38H and S449* mutations, respectively. The pET15b/*cj1411c* vector was used as a template for mutagenesis PCR, using a Quikchange II kit as described (Section 2.24). With the exception of the S449* mutant, the mutated gene constructs were digested with NdeI to excise the unwanted N-terminal region. The K38H and E109H mutant plasmids were then run on agarose gels to separate the cleaved sequence from the rest of the plasmid (Figure 4.20). The E109H NdeI digest produced a small DNA fragment at the expected size (327 bp). A very faint band was detected (114 bp) following NdeI digestion of the K38H mutant, which was again at the expected size. The pET15b/mutant gene constructs were gel extracted using a Qiagen kit, and the DNA recircularised using a quick ligase kit (NEB). The plasmids were transformed into competent *E. coli* XL1 Blue cells and grown on selective LB agar. Plasmids were recovered from successful transformants and the sequences of the mutant plasmids checked via DNA sequencing. The S449* mutant plasmids were also sent off for sequencing to check for insertion of the stop codon before being used in protein expression trials.
Figure 4.20. NdeI digestion of mutagenised pET15b/cj1411c. The agarose gel (0.8 %) shows the NdeI restriction digests products (C, double lanes) of mutants K38H and E109H and the non digested plasmid (U, single lanes). The digested K38H produced two bands at ~6.9 kbp and 0.1 kbp. The E109H produced two bands at ~6.7 kbp and 0.3 kbp. A 2-log DNA ladder was loaded to determine band sizes.
4.15.1. Expression trials of truncated versions of CYP172A1

The truncated versions of the cj1411c gene were transformed into HMS174 (DE3) cells, and the growth and expression conditions that were used for the full length protein were again applied (Section 2.29). Small scale grow ups were initially used to check that the truncated proteins were soluble. Following a 16 hour induction period, samples were removed and treated with a BugBuster solution in order to determine total, soluble and insoluble forms of the truncated proteins (Figure 4.21). Apart from E109H, the majority of the truncated versions of CYP172A1 appeared to be insoluble. The western blot of the gel confirmed that the new sizes of the truncated proteins were ~50 kDa (K38H), ~41 kDa (E109H) and ~54 kDa (S449*).

Large scale protein expression of the K38H, E109H and S449* truncated mutants was undertaken to obtain enough protein for subsequent experiments. Only the E109H truncation mutant was able to produce a soluble red protein. Even with the addition of δ-ALA to all the flasks, no red (heme-containing) protein was produced for the other two CYP172A1 truncations. All the samples were subjected to Ni-NTA affinity purification in order to detect His-tagged protein production. Samples taken at specific points during the purification process were separated by SDS-PAGE and then blotted onto a PDVF membrane. The membranes were probed with an anti-His antibody to determine protein expression. Bands that corresponded with the correct sizes for the K38H and S449* truncations were observed upon Western blotting (Appendix, Figure S10). For the K38H truncation these fractions were eluted from the Ni-NTA resin at 60 mM imidazole, and for the S449* truncation with 30 mM imidazole. The fractions that contained anti-His tag reactive bands were analysed by UV-Visible spectroscopy, but no Soret peak was observed at 418 nm for either truncated protein. This suggested that heme incorporation had not occurred. The C-terminal truncation (S449*) might have caused problems for heme incorporation, as the highly conserved heme binding motif (FxxGxxxCxC) is located towards the Cterminal end of the protein (45 amino acids away from the introduced stop codon). The reason for the lack of heme incorporation for the K38H is less certain, and without structural detail for this P450 it is difficult to identify with certainty the reason for the failure of the mutant to incorporate heme.
Figure 4.21. SDS-PAGE and Western Blot of truncated CYP172A1 mutants. Expression cell samples from K38H, E109H and S449* truncation mutants were treated with BugBuster to examine the total protein expressed (TP) and the proteins in the soluble (S) and insoluble fractions (INS). The major band of the K38H CYP172A1 has an apparent $M_w$ of $\sim$50 kDa, whereas the E109H truncation mutant has a $M_w$ of $\sim$41 kDa and the S449* truncation mutant has a $M_w$ of $\sim$54 kDa, which are all consistent with the predicted sizes of the truncated version of CYP172A1. Full length CYP172A1 (54.4 kDa) was loaded as a control (C). Samples were run on a 12 % SDS-PAGE gel and bands compared to a protein marker (M).
As mentioned above, expression and recovery of the E109H CYP172A1 truncation mutant was achieved (Figure 4.22). After Ni-NTA purification (as for full length CYP172A1) the protein was buffered exchanged into a 100 mM Tris buffer, with 250 mM KCl at pH 8.0. However, this change in buffer caused the protein to precipitate, which was unexpected as the full length protein did not exhibit the same characteristics. As a result, the E109H truncation mutant P450 was kept in a 100 mM KP, buffer, with 100 mM KCl at pH 7.5 for further purification in order to prevent protein precipitation. UV-Visible spectroscopy of this mutant produced a spectrum atypical of a P450. The heme Soret peak was observed at 420 nm, which might suggest some contamination from imidazole, despite several dialysis steps being used. The Soret peak was not as intense a feature as observed for the full length CYP172A1, and the 418:280 ratio (Rz value) was never above 0.5 (compared to 1.3 in intact CYP172A1), which suggested that heme incorporation was not complete. The protein was then subjected to size exclusion chromatography (Tricorn Superdex 10/30), and selected fractions were analysed by SDS-PAGE (Figure 4.22).

Whilst a dominant protein band was observed at the correct M_w (~41 kDa), bands of higher molecular mass were also observed on the gel (Figure 4.22). Based on the broad elution profile from the size exclusion column, it is highly likely that these bands are higher order aggregates, as seen in the full length CYP172A1. Because of the apparent incomplete heme incorporation and the presence of higher aggregates, the E109H truncation mutant, like the K38H and S449* CYP172A1 truncations, was not studied further.
Figure 4.22. CYP172A1 E109H purification using size exclusion gel chromatography. A) Elution profile of E109H Passed through a size exclusion (S-200) gel chromatography column. Two wavelengths were measured, 280 nm (blue line) and 420 nm (red line). Protein elution occurred between 8 ml (void volume) and 20 ml. B) 12 % SDS-PAGE gel analysis of S200 fractions B11-C4. A dominant band was observed at ~41 kDa (the expected size for the E109H truncation mutant) using the protein marker (M) as guide. Aggregated protein likely explains a number of the protein bands observed at higher M_w’s.
4.16. Structure determination of CYP172A1

The amino acid sequence of CYP172A1 is well conserved among C. jejuni strains, but the amino acid sequence identity to other P450 structures that have been solved by crystallography was at best 28\% (the A328V mutant of the P450 BM3 heme domain PDB, 1JPZ). This makes CYP172A1 an orphan P450 enzyme (in terms of uncertainty regarding its substrate identity) with no similar protein structure currently known. For this reason it was important to try and obtain a crystal structure in order to define how this unusual P450 folds. Crystallography was performed in 96 well plates using commercial screens and the full length CYP172A1, which was prepared both with and without 100 mM BME. Despite the wide range of conditions investigated, no crystals were formed. The most probable reason was that aggregation of CYP172A1 occurred and that no single monodispersed species formed. The protein prepared with 100 mM BME failed to produce crystal structures, possibly in part due to BME oxidising over the time required for crystal growth. The addition of saturating levels of 1-PIM was also trialled in attempts to stabilise CYP172A1. However, no crystals were formed in this case either. In the majority of conditions tested the protein either precipitated or failed to form a crystal-like growth (Figure 4.23). For the 1-PIM bound protein, non-regular crystalline material was observed. However, these objects were not true crystals, but instead structures termed “sea urchins”, which are aggregated protein and which are not conducive to crystal formation. [264]

As crystal trials proved unsuccessful, protein modelling was instead used to gain an understanding of the topology of CYP172A1. As previously mentioned, there are no crystal structures of P450s strongly related to CYP172A1. The protein sequence was entered into the PDB databank (May 2013) to identify the most related protein. There were a number of candidate proteins, including many P450 BM3 mutants with the best similarity being the 1JPZ structure for the N-palmitoylglycine-bound BM3 heme domain (with 28\% amino acid sequence identity to CYP172A1). This solved structure was used as a template to produce a model structure for CYP172A1 (Figure 4.24). The region of highest sequence similarity was towards the C terminus, especially around the amino acids involved in heme binding, which was not surprising as many P450s display high levels of amino acid similarity in this region (Figure 1.11). From the model we can
see that the heme (red) is located in the centre of the enzyme with the distal water ligand present (pink). The I helix sits above the heme and several other helices (coloured cyan) make up the majority of the secondary structure in the model. Two out of the three cysteine residues in CYP172A1 are shown in spacefill. The third is located at the very close to the N terminus (Cys4), and does appear in the homology model. The positions and surface exposure of the side chains of the two non-heme ligand cysteine residues (Cys204 and Cys4) are at present not known, but based on BME experiments it is possible that they might be solvent accessible and could contribute to inter-P450 disulfide bonds. The final cysteine is the heme iron-ligating Cys399, shown below the plane of the heme on the opposite face from the I helix.

![Figure 4.24. Representation of a homology model structure of CYP172A1](image)

A homology model was constructed based using the PDB P450 structure 1JPZ as a template. The β sheets are shown as magenta arrows, with α helices as cyan ribbons and the heme (red sticks) in the middle of the model with its Fe atom (chocolate) at its centre. Two of the three cysteine residues are also highlighted in atom coloured space fill. Cys 399 is phylogenetically conserved and coordinates to the heme iron. The second cysteine (Cys204) appears on the G α helix.
Figure 4.23. Examination of crystalline material from trials of CYP172A1 in presence and absence of 1-PIM. A) Phase separation seen for CYP172A1 (0.1 M ammonium acetate, 0.1 M Bis Tris buffer, pH 5.5, 17 % w/v PEG 10K ), B) and F) CYP172A1 precipitation (0.2 M zinc acetate, 0.1 M sodium acetate buffer, pH 4.5, 10 % w/v PEG 3K and 24 % w/v PEG 1500/20 % v/v glycerol, respectively). C) Protein remaining in solution (0.1 M Bis Tris buffer, pH 5.5 with 3.0 M sodium chloride). D) and E) “Sea urchin” formation with 1-PIM bound in 0.2 M zinc acetate, 0.1 M imidazole pH 8.0, 20 % w/v PEG 3K).
4.17. Summary

The novel P450, CYP172A1, from *C. jejuni* was expressed and purified using a recombinant *E. coli* expression system. Following extensive trials, by varying the experimental conditions, the *E. coli* HMS174 (DE3) cell line was chosen as the most suited for the production of soluble protein (Figure 4.4), which was possibly due to the engineering of this strain for efficient codon usage to suit *cj1411c* gene expression. With the soluble protein, general P450 characterisation was undertaken, including the binding of CO to the ferrous heme iron in order to form the P450 complex. This was first identified as a unique feature of the P450 class of proteins by Omura and Sato in 1964 [186]. In the presence of sodium dithionite, the reduced CYP172A1 readily formed the P450 complex (even in the presence of oxygen) with a small amount of P420 complex formation. The P420 state is caused by the loss of the cysteine thiolate linkage to the heme, as the cysteinate is protonated to the thiol in this form [239]. The P450 was also able to form an NO complex, with a red shift of the Soret peak to 432 nm, similar to that reported for other P450s. EPR studies confirmed that, under native conditions, the ferric protein is in a LS state, and that the heme iron is coordinated to the protein via a cysteinate linkage.

The molar extinction coefficient for the CYP172A1 ferric, LS heme was calculated to be $\varepsilon_{418} = 105 \pm 5 \text{ mM}^{-1} \text{ cm}^{-1}$ and the coefficient was used to calculate the concentration of any CYP172A1 protein sample. This value was very similar to the value reported for P450 BM3 [242]. Far UV CD showed that, under resting conditions, the protein is predominantly alpha helical, as is known to be the case for all known P450s. The secondary structure was almost completely lost following the addition of 2.25 M GdnHCl. The stability of CYP172A1 tertiary structure towards GdnHCl, as measured using fluorescence spectroscopy, was quite low by comparison to thermostable CYP175A1, but similar to the *B. subtilis* P450 BioI, suggesting a typical level of structural stability for a mesophilic P450 [143, 249].

Perhaps the most interesting result from this chapter is the discovery of the tendency of CYP172A1 to self-aggregate, forming oligomeric/multimeric states, as visualized by MALLS, SDS-PAGE and size exclusion chromatography. The majority of bacterial
P450s are considered to be monomeric, and many that have been crystallized are monodisperse. CYP172A1 is clearly polydisperse and analysis with MALLS produced results that indicated the P450 formed higher order structures that ranged from tetramers up to 18-mers, with a range of sizes in between. Using a size exclusion column the majority of CYP172A1 eluted with the void volume, and later eluting protein produced a broad peak. On SDS-PAGE separation, higher molecular weight bands were still visible, and were also detectable using mass spectrometry and immunoblotting, even with the presence of strong reducing conditions and with a five minute boiling period at 95 °C in attempts to disaggregate the protein.

The origin of the interactions that promote CYP172A1 self-aggregation is not fully clear, but such protein-protein interactions can be facilitated by Van der Waals, hydrogen bonding, ionic interactions or disulfide bridge formation. The first three of these types of interactions are generally considered to be relatively weak. However, collectively they can combine to produce a strong inter-molecular attraction. Disulphide bridges would not have been immediately expected for CYP172A1, as there are only three cysteine residues in the P450 (Cys4 Cys204 and Cys399), the last of which is the conserved cysteine residue that coordinates the heme iron, and so is almost certainly unable to form any disulfide bridges with a cysteine from another CYP172A1. The addition of 100 mM BME to CYP172A1 indicated some disaggregation of CYP172A1 (although a mixture of oligomeric forms is still present), and suggests that protein aggregation may be partially stabilized by disulfide bridges, but that it is not the only form of molecular interaction underlying CYP172A1 oligomer formation.

In order to address issues associated with the aggregation of CYP172A1, several techniques were used in efforts to produce protein that was less polydisperse and therefore more amenable to crystallization. Firstly, a range of different buffer and additives were screened using the thermofluor assay to explore buffer conditions that stabilize the P450 [257]. In the initial purification buffer (100 mM Tris, with 100 mM KCl at pH 7.5) the protein was stable up to 55 °C, but with an alteration in the buffer composition to KP, and with the addition of glycerol, the T_m was raised by 2.7 °C. However, despite the change in buffer, CYP172A1 was still found to oligomerise.
Purification tags are known to cause protein aggregation in other proteins [265]. Cleavage of the His<sub>6</sub> tag was successfully achieved using thrombin. However, analysis of the resulting fractions indicated that the removal of the tag did not decrease the extent of the protein aggregation, and thus that the tag was not a major cause of the CYP172A1 aggregation.

Truncations of the P450 were used to see if flexible parts at the N- or C-terminus of the protein might act to nucleate CYP172A1 for aggregation. Three truncations were produced; two from the N-terminus and one at the C-terminus, namely K38H, E109H and S449*, respectively. Two truncations (K38H and S449*) caused the failure to incorporate heme into the P450, and the heme binding was also severely disrupted in the E109H truncation mutant. Purification of the E109H truncation of CYP172A1 was achieved. However, the presence of higher M<sub>w</sub> bands (observed in SDS-PAGE studies) suggested that protein aggregation still occurred and that this truncation did not prevent P450 interactions to form oligomers.

Crystallography of proteins usually requires that a species is monodisperse so that a regular crystal structure can form. CYP172A1 displays polydispersity and it is clear that multiple higher order forms of the P450 exist, and that some are likely in equilibrium. Because of this property, it was not expected that crystallography would be successful for CYP172A1. A wide range of commercial screens were used to try and find any condition that would promote CYP172A1 crystal formation. Over 90% of the screening conditions tested produced protein precipitants within 24 hours of CYP172A1 addition. No improvement was observed with the addition of the inhibitor 1-PIM, despite its stabilizing effect, leading to an increase in the T<sub>m</sub> of CYP172A1 to almost 60 °C. With the inability to produce CYP172A1 crystals, homology modeling was instead used to gain an appreciation of the structure of CYP172A1. The main issue with performing this type of modeling study was that CYP172A1 is currently an orphan protein (with respect to substrate specificity) and that all reported P450 structures have protein sequences quite dissimilar to that of CYP172A1. However, the P450 structures that have been solved to date illustrate that the overall P450 fold is well conserved, with 26 unique structures of different P450s accounting for almost all the 600 solved P450 structures in
the PDB [124]. The model generated for CYP172A1 illustrated that the protein was
predominantly alpha helical, in agreement with the data obtained from the CD
experiments, and as expected for a P450 enzyme. However, few other conclusions can
be drawn on the accuracy of the model or its predictive capacity at the current time.

This chapter has focused on examination of the biophysical aspects of the sole P450
found in C. jejuni, CYP172A1. Building on this work, Chapter 5 will examine the
P450’s ability to interact with ligands, in order to determine what type of molecules act
as either substrate- or inhibitor-like compounds.
Chapter 5. Ligand binding to CYP172A1

5.1. Introduction

A further area of interest into the investigation of the novel CYP172A1 was to identify substrates and inhibitors of the P450, which might help to establish the role of the P450 in its natural host, *C. jejuni*. There are no reports of type I (substrate-like) optical binding to CYP172A1, and there are no strongly related protein structures solved with their respective substrates bound, as mentioned in the previous chapter, that might be used as a guide to CYP172A1 substrate selectivity. This meant that the substrate could be any number of molecules, e.g. lipid, sugar, polyketide etc, and so the identification of the preferred substrate would be important in identifying the role of CYP172A1 in *C. jejuni*. Like most enzymes, P450s have inhibitors as well as substrates. Fortunately the identification of each molecule type can often be determined from the induced shift in the Soret peak as either a red shift or a blue shift for inhibitor or substrate, respectively. Without a structure of the enzyme and, in particular, prior knowledge of the composition of the active site, the binding affinities of different compounds determined experimentally can help to give clues as to the topology of the catalytic centre of the enzyme. The tighter the calculated dissociation constant the more closely that substrate-like molecule will likely match the natural substrate in the host cell. The structure of the substrate could then be used as a model to look for likely natural candidate substrate molecules.

With the identification of a substrate, it could then become possible to perform more elaborate studies, such as analysing the effect that the substrate has on the hemoprotein’s redox potential (achieved by using redox potentiometry). Typically the binding of a substrate will shift the midpoint potential to a more positive value to allow electrons to flow to the P450 heme from the redox partner. The control of the redox potential serves as an important regulatory step for P450 enzymes. By performing enzymatic turnover (TO) reactions, an appreciation for the type of catalysis that CYP172A1 is able to perform can again lead to understanding the reactions that the enzyme may conduct within the host, possibly at the CPS. As TO reactions require
redox partners these experiments can provide information about the preferred redox partner systems that the P450 would use to perform catalysis in its host C. jejuni. C. jejuni does contain both flavodoxin and ferredoxin proteins and thus by measuring the electron flow with non native redox partners that include ferredoxin and flavodoxin components, an indication of the preferred redox system can be suggested.

The aim of this chapter is thus to study interactions of inhibitor and substrate/substrate-like molecules with CYP172A1 in order to gain information on the nature of the active site within CYP172A1, and to describe the molecular shape and characterisation of the compounds that can bind to this P450. In the absence of a crystal structure, this was considered to be a good route to gaining an appreciation for the type of molecules CYP172A1 may bind in C. jejuni. By performing CYP172A1 mediated TO reactions using substrate-like molecules with heterologous redox partner systems, then the catalytic capability of the protein can also be determined and the type of reaction elucidated.

5.2 Type II binding molecules
As mentioned in the opening chapter (Section 1.6) type II (inhibitor) binding compounds exhibit particular phenomena that are readily detected by UV-Visible spectroscopy. The most prominent feature being the red shift of the Soret peak to longer wavelengths. This occurs when the inhibitor coordinates to the ferric heme iron at the distal position (occupied by a water molecule under resting conditions). The nature of the interaction will affect the degree of the shift, nitrogen ligated inhibitors (e.g. azoles) cause a shift to ~425 nm, nitric oxide to ~430-435 nm (for ferric P450s) and CO to ~450 nm (for ferrous P450s) [239]. The differences in the shifts are related to the field strengths, whereby a species with a high field strength produces a greater shift in the Soret peak [266]. The binding of inhibitors typically makes the P450 enter a LS state and an increase in the intensity of the heme δ band (~360 nm) is often observed alongside changes to the Soret features.
5.2.1 Azole binding to P450s

Azoles, which were originally developed for their antifungal properties, are well known inhibitors of P450s. They are largely hydrophobic, 5-membered heterocyclic ring compounds which helps them form favourable interactions within the P450 active site. Coordination to the heme iron occurs through a nucleophilic nitrogen atom from the azole moiety, which is usually an imidazole or triazole based derivative. Optical titrations were carried out using a range of azoles, including econazole, miconazole and ketoconazole, and in total the binding of six azoles was assayed. All produced type II optical shifts, but the degree of the shift was different. In tighter binding azoles the change in the Soret peak was greater than those of the weaker binding ones.

Econazole was the tightest binding inhibitor and produced a spectral shift in the Soret peak from 418 nm to 423 nm (Figure 5.1) as well as an increase in peak intensity at 360 nm (δ band) indicating an increase in the LS state. In the Q-band region, there was a decrease in peak intensity at the α band (585 nm) and the β band shifted slightly to a longer wavelength. The difference plot, which was calculated by subtracting the ligand-free spectrum from each successive ligand-bound spectrum in the set, demonstrated that as more econazole was added the trough at 417 nm deepened and that there was an increase in the peak with an absorbance at 436 nm, typical for an inhibitor binding difference spectrum. Taking the difference in the peak and trough at these two wavelengths at each concentration, it was possible to generate a binding curve (Figure 5.1), which was used to calculate a dissociation constant ($K_d$) value of $0.10 \pm 0.01 \mu M$. This nanomolar binding affinity was comparable to $K_d$’s reported for other P450s with econazole [267].
Figure 5.1. Optical titration of CYP172A1 with econazole. A) Selected UV-Vis absorption spectra of CYP172A1 (3.3 μM) titrated with econazole. Additions (0.1 μl) of an econazole stock (1 mM) were added to a final concentration of 4 μM. The resting Soret peak was at 418 nm and decreased in intensity and shifted to 423 nm with the addition of econazole. Insert, difference spectra which were used to determine the maximum absorption changes upon ligand binding. B) Binding curve from the titration, calculated by subtracting absorption values at the trough (417 nm) from those at the peak (436 nm) in each difference spectrum in order to obtain overall absorbance change values at each [econazole]. The difference values were then plotted against [econazole] and the data were fitted using equation 2 to generate a $K_d$ of 0.10 ± 0.01 μM.
Titrations with miconazole produced a typical type II shift where the Soret peak moved from 418 nm to 423 nm. The absorption intensity from the α band decreased as more miconazole was added but the δ band increased in intensity, again indicating the protein was in a LS state (Figure 5.2). From the difference spectra a peak and trough were observed at 435 nm and 414 nm, respectively. The difference in the values between these maxima and minima were plotted against the concentration of miconazole added in order to generate a binding curve which was used to determine a $K_d$ value of $0.17 \pm 0.05 \mu M$. This value was very similar to that achieved for econazole, which was not surprising as the two compounds have very similar structures (Figure 5.3). The dissociation constant for miconazole may have been slightly higher due to the extra chlorine atom which may not be favoured in the lipophilic active site. However, the values were almost identical within error.

Figure 5.3. Chemical structures of A) econazole and B) miconazole. Both structures contain a nucleophilic nitrogen which coordinates to the heme iron and two chlorine substituted phenolic rings.
Figure 5.2. Optical titration of CYP172A1 with miconazole. **A)** Selected UV-Vis absorption spectra of CYP172A1 (3.8 µM) titrated with miconazole. Titrations were done with a 1 mM stock, as described in Figure 5.1. The Soret peak shifted from 418 to 422 nm with addition of miconazole. Insert shows difference spectra used to determine maximum absorption changes on ligand binding. **B)** Binding curve from the titration. Values were calculated by subtracting $A_{414}$ (trough) from $A_{435}$ (peak) for each [miconazole], and data plotted against [miconazole] and fitted using equation 2 to generate a $K_d$ of 0.17 ± 0.05 µM.
Clotrimazole also produced a type II spectral shift when titrated against CYP172A1. However, unlike in the cases of miconazole and econazole the changes associated with binding were not as strong. Firstly the Soret peak shifts to a lesser extent (418 nm to 421 nm) and the changes to the α band were much less notable (Figure 5.4). The δ band appeared to increase in intensity. However, binding to CYP172A1 may not have been the only reason for the smaller heme absorption changes. The solubility of clotrimazole was poor in aqueous solutions and with each addition the turbidity of the solution increased. Evidence for this was observed at 280 nm, in the final spectrum (purple) a large increase in absorbance was observed. Despite this, it was still possible to generate a binding curve (Figure 5.4) which was used to estimate a $K_d$ of $2.81 \pm 0.11 \mu M$. This value was considerably weaker compared to the $K_d$ for clotrimazole with C. albicans and human CYP51 [268], which suggest that the active site topology was very different in these two different classes of P450 enzymes.

The calculated $K_d$’s for all of the azoles titrated against CYP172A1 were compared (Table 5.1). From the six azoles tested the best inhibitor was econazole, which was closely followed by miconazole and then ketoconazole. Econazole and miconazole had structural similarity, having a short alkyl chain (three carbons) bridged by an ether bond, and the carbon attached to a chlorinated benzene ring at either end (Figure 5.3). The central carbon is bonded to an imidazole nitrogen, with the other nitrogen atom on this ring being responsible for coordinating to the heme iron. The structure of ketoconazole is much larger than the previous two azoles and it might be that it extends out of the active site, but in such a way that it makes favourable contacts along the protein to enable it to have a low $K_d$. The structures of the other three azoles that bound with micromolar affinities have bulkier groups that are more closely arranged, and it is possible that they suffer, to some extent, from steric hindrance with amino acid side chains within the enzyme’s active site. The smaller shifts in the Soret peak for some of these azoles might also indicate that the water which occupies the distal ligand position on the heme was not being removed on their binding, as typically a shift to 424-425 nm is observed with direct imidazole nitrogen coordination to the iron atom [239]. It is possible the nitrogen coordinates to the heme through the distal water molecule that is retained on the heme iron. This a type of binding that is becoming more recognised in
The triazole based azoles (fluconazole and voriconazole) have weaker dissociation constants for CYP172A1, which has also been reported for CYP51 and might result in part from the poorer ligand field strength compared to the imidazole based azoles [270].

<table>
<thead>
<tr>
<th>Azole</th>
<th>Soret peak shift</th>
<th>Calculated $K_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Econazole</td>
<td>418 - 423</td>
<td>100 ± 10 nM</td>
</tr>
<tr>
<td>Miconazole</td>
<td>418 - 422</td>
<td>170 ± 50 nM</td>
</tr>
<tr>
<td>Ketoconazole</td>
<td>418 - 421</td>
<td>508 ± 121 nM</td>
</tr>
<tr>
<td>Clotrimazole</td>
<td>418 - 421</td>
<td>2.81 ± 0.11 μM</td>
</tr>
<tr>
<td>Voriconazole</td>
<td>418 - 421</td>
<td>5.8 ± 0.2 μM</td>
</tr>
<tr>
<td>Fluconazole</td>
<td>418 - 421</td>
<td>20 ± 2 μM</td>
</tr>
</tbody>
</table>

**Table 5.1. Summary of azole binding to CYP172A1.** The table details the azole-induced changes to the Soret peak position as well as the calculated $K_d$ for each inhibitor used.
Figure 5.4. Optical titration of CYP172A1 with clotrimazole. A) Selected UV-Vis absorption spectra of CYP172A1 (3.9 µM) titrated with clotrimazole. Titrations were done with a 10 mM stock, as described in Figure 5.1. The Soret peak shifted from 418 to 421 nm with addition of clotrimazole. Insert shows difference spectra used to determine maximum absorption changes on ligand binding. B) Binding curve from the titration. Values were calculated by subtracting \( A_{414} \) (trough) from \( A_{432} \) (peak) for each [clotrimazole], and data plotted against [clotrimazole] and fitted using equation 2 to generate \( K_d \) of 2.81 ± 0.46 µM.
5.2.2. Further type II binding
Along with the azoles, other type II inhibitor molecules are known, such as cyanide and imidazole. Cyanide is a potent poison to humans as it binds to cytochrome c oxidase preventing it from functioning as part of the electron transport chain [271]. Cyanide is a small ionic ligand and its binding produces a distinctive spectral shift due to the coordination of its carbon atom to the heme iron (Figure 5.5). Upon binding, the Soret peak at 418 nm decreased in intensity as the water was displaced by the cyanide, and the new Soret feature emerged at 435 nm. The δ band also increased in intensity which, together with the rising peak at 435 nm, indicated that cyanide made the P450 heme iron more LS. The α and β bands also lose their separate definition as they merge into a single peak at ~560 nm due to the changing heme environment. A dissociation constant for the binding of (sodium) cyanide was calculated from a binding curve as 1.45 ± 0.39 mM. This value was very similar to that reported for P450 BM3 (1.68 mM) [272] but significantly tighter than the observed $K_d$ for the P450-phthalate dioxygenase reductase fusion enzyme CYP116B1 (9.7 ± 0.6 mM) [273], which highlights the variability in the binding of cyanide to P450s. Cyanide is frequently reported to be a weak inhibitor for P450s (with a dissociation constant in the millimolar range), which is probably due to its charge being unfavourable in a lipophilic environment and also the increase in electron density placed on the heme iron [274].

A concentrated stock of imidazole (1 M) was prepared to perform an optical titration, as for many P450s imidazole does not have a strong binding affinity. During the titration with CYP172A1, the starting Soret peak (418 nm) decreased and a peak began to emerge at 425 nm. However, a clear conversion to a peak distinct from the ferric resting state was not observed (Figure 5.6). The δ band also increased in absorbance and the α band absorption peak decreased in intensity as more ligand was added. The imidazole reinforced the LS state of CYP172A1, and it can be concluded that it bound in a type II manner. A binding curve generated from the maxima and minima taken from a set of difference spectra (435 and 416 nm, respectively) was used to estimate the $K_d$ as 3.28 ± 0.07 mM. This makes imidazole a weaker binding inhibitor compared to azole drugs. This result highlights the importance of the large hydrophobic side chains in molecules such as econazole in making for a tight binding molecule to P450s. The $K_d$ value for
imidazole was different to some others reported, being tighter than the $K_d$ for CYP121 (64 mM), but weaker than that reported for the RDX-degrading XplA (160 µM) [267, 275]. As with cyanide, imidazole is a polar molecule and this charge is disfavoured in a hydrophobic environment, which explains the weak binding. The active site of XplA is relatively polar, hence the tighter binding observed with imidazole.
Figure 5.5. Optical titration of CYP172A1 with sodium cyanide. A) Selected UV-Vis absorption spectra of CYP172A1 (16 µM) titrated with sodium cyanide. Additions (1 µl) of a sodium cyanide stock (2M) were made to a final concentration of 60 mM. The resting Soret peak was at 419 nm and decreased in intensity and shifted to 439 nm with addition of the sodium cyanide. Insert, the difference spectra which were used to determine the maximal absorption changes upon ligand binding. B) Binding curve from the titration, calculated by subtracting absorption values at the trough (419 nm) from those at the peak (446 nm) in each difference spectrum in order to obtain total absorption change values at each [cyanide]. The difference values were then plotted against [cyanide] and the data fitted using a sigmoidal function to generate a $K_d$ of 1.45 ± 0.39 mM.
Figure 5.6. Optical titration of CYP172A1 with imidazole. A) Selected UV-Vis absorption spectra of CYP172A1 (2.5 µM) titrated with imidazole. Titrations were done with a 1 M stock, as described in Fig. 5.1. The Soret peak shifted from 417 to 437 nm with addition of imidazole. Insert shows difference spectra used to determine maximum absorption changes on ligand binding. B) Binding curve from the titration. Values were calculated by subtracting A_{417} (trough) from A_{437} (peak) for each [imidazole], and data plotted against [imidazole] and fitted using equation 2 to generate a $K_d$ of 3.28 ± 0.07 mM.
5.2.2.1 Binding of phenyl imidazoles (PIMs) to CYP172A1

To further investigate the role of additional components to the core imidazole structure to improve binding to CYP172A1, PIMs were titrated against the P450. They contain an addition phenyl group that is attached to the imidazole at the 1, 2 or 4 position (Figure 5.7). Titrations with 1-PIM resulted in an increased heme absorbance in the δ band and the Soret peak shifting from 418 to 422 nm (Figure 5.8). A binding curve was generated using the changes in the peak and trough from the difference spectra (434 and 415 nm, respectively), with absorption difference plotted against the ligand concentration and a dissociation constant of $38.67 \pm 1.67 \mu M$ calculated. A very similar result emerged when titrating with 4-PIM (Figure 5.9). As with the case for 1-PIM, there was a slight shift of the Soret peak, the α band signal was reduced and the δ band increased with the addition of 4-PIM. The difference spectra looked almost identical to those for 1-PIM and a very similar $K_d$ of $39.77 \pm 1.88 \mu M$ was calculated. This value was almost identical as the value reported for P450$_{cam}$, but much tighter than the value observed for CYP158A2. These results reinforce the importance of the active site composition in the binding of compounds inhibitor compounds to P450s [244, 276].

![Chemical structures of 1-, 2- and 4-PIMs](image)

**Figure 5.7. Chemical structures of 1-, 2- and 4-PIMs.** The phenyl ring is bound to the imidazole ring at 3 different positions.

There was no detectable binding observed with 2-PIM (i.e. from heme coordination). This might suggest that there was a steric clash preventing the molecule from positioning the available nitrogen above the heme iron. Alternatively, 2-PIM may occupy a different binding mode in the active site which did not involve iron coordination. Both 4-PIM and 1-PIM had binding affinities 100 times tighter for CYP172A1 compared to that for imidazole. This was likely due to the phenyl ring forming favourable interactions with one or more active site amino acid side chains.
Figure 5.8. Optical titration of CYP172A1 with 1-PIM. A) Selected UV-Vis absorption spectra of CYP172A1 (2.5 µM) titrated with 1-PIM. Titrations were done with a 0.8 mM stock, as described in Figure 5.1. The Soret peak shifted from 418 to 423 nm with addition of 1-PIM. Insert shows difference spectra used to determine maximum absorption changes on ligand binding. B) Binding curve from the titration. Values were calculated by subtracting $A_{415}$ (trough) from $A_{434}$ (peak) for each [1-PIM], and data plotted against [1-PIM] and fitted using equation 2 to generate a $K_d$ of 38.67 ± 1.46 µM.
Figure 5.9. Optical titration of CYP172A1 with 4-PIM. A) Selected UV-Vis absorption spectra of CYP172A1 (2.5 µM) titrated with 4-PIM. Titrations were done with a 1 mM stock, as described in Figure 5.1. The Soret peak shifted from 418 to 423 nm with addition of 4-PIM. Insert shows difference spectra used to determine maximum absorption changes on ligand binding. B) Binding curve from the titration. Values were calculated by subtracting $A_{414}$ (trough) from $A_{434}$ (peak) for each [4-PIM], and data plotted against [4-PIM] and fitted using equation 2 to generate a $K_d$ of 39.77 ± 1.46 µM.
5.2.2 Binding of fatty acid linked imidazoles to CYP172A1

Another set of inhibitor molecules are fatty acid linked imidazole compounds (Im-C10, -C11 and -C12, indicating imidazole linked to the ω-carbon of saturated fatty acids with chain length C10-C12) [277]. These compounds were titrated against CYP172A1 and type II shifts were observed (Appendix Figure S11), with the Soret peak moving to a final position of 423 nm. From binding curves the $K_d$’s were calculated as, Im-C10 = 216.0 ± 8.1 µM, Im-C11 = 148.8 ± 8.3 µM and Im-C12 = 51.0 ± 4.4 µM. The increase in fatty acid chain length improved the affinity of binding, likely due to the alkyl chain forming more hydrophobic interactions within the active site. In this way, addition of methyl groups may result in the longer molecules making more favourable interactions than the shorter ones. The dissociation constants of these molecules for CYP172A1 were rather weak when compared to P450 BioI (2.6, 1.5 and 1.0 µM for Im-C10, Im-C11 and Im-C12, respectively) and 10 times weaker than the dissociation constants obtained for P450 BM3 [143, 277]. At higher ligand concentrations, turbidity proved to be an issue due to the limited solubility of these fatty acid linked imidazoles in a predominantly aqueous environment, and increasing background absorbance was observed.

Several different type II ligands were identified for CYP172A1, which exhibited a range of dissociation constants and shifts to the Soret peak (Table 5.2.). The tightest binding compounds were the azoles, and especially econazole and miconazole. The polar molecules, cyanide and imidazole bound less tightly, likely due to their incompatibility with the apolar CYP172A1 active site. The addition of fatty acids or phenyl groups improved tightness of binding for imidazole based molecules to the P450.
Chapter 5. Ligand binding to CYP172A1

<table>
<thead>
<tr>
<th>Compound</th>
<th>Soret Shift (nm)</th>
<th>$K_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO</td>
<td>418-448</td>
<td>ND</td>
</tr>
<tr>
<td>Cyanide</td>
<td>419-439</td>
<td>1.45 ± 0.39 mM</td>
</tr>
<tr>
<td>NO</td>
<td>418-432</td>
<td>ND</td>
</tr>
<tr>
<td>Clotrimazole</td>
<td>418-421</td>
<td>2.81 ± 0.11 µM</td>
</tr>
<tr>
<td>Econazole</td>
<td>418-423</td>
<td>0.01 ± 0.01 µM</td>
</tr>
<tr>
<td>Fluconazole</td>
<td>418-421</td>
<td>20 ± 2 µM</td>
</tr>
<tr>
<td>Ketoconazole</td>
<td>418-422</td>
<td>0.508 ± 0.121 µM</td>
</tr>
<tr>
<td>Miconazole</td>
<td>418-423</td>
<td>0.17 ± 0.05 µM</td>
</tr>
<tr>
<td>Voriconazole</td>
<td>418-421</td>
<td>5.8 ± 0.2 µM</td>
</tr>
<tr>
<td>Imidazole</td>
<td>419-425</td>
<td>3.28 ± 0.07 mM</td>
</tr>
<tr>
<td>1-PIM</td>
<td>419-423</td>
<td>38.67 ± 1.46 µM</td>
</tr>
<tr>
<td>2-PIM</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4-PIM</td>
<td>419-423</td>
<td>39.77 ± 1.78 µM</td>
</tr>
<tr>
<td>Im-C10</td>
<td>419-423</td>
<td>216.0 ± 8.1 µM</td>
</tr>
<tr>
<td>Im-C11</td>
<td>419-422</td>
<td>148.8 ± 8.3 µM</td>
</tr>
<tr>
<td>Im-C12</td>
<td>419-422</td>
<td>51.0 ± 4.4 µM</td>
</tr>
</tbody>
</table>

Table 5.2. Summary of type II compounds binding to CYP172A1. The shifts in the Soret peak along with the calculated $K_d$ for a range of type II compounds are listed. ND; not determined, in the case of CO and NO, N/A; not available as heme interaction did not occur.

5.3. Type I ligands

Type I ligands occupy the P450 active site in a different way to inhibitors. They bind in the active site pocket and rather than interacting directly with the heme iron (as in the case for type II ligands), they generally act as substrates. Binding of the substrate results in the release of the axial water molecule resulting in the formation of a 5-coordinate state and a Soret peak shift to a lower wavelength. This also facilitates the transition of the heme iron from LS to a HS state. One report, which compared the CPS from wild type C. jejuni and a cj1411c knockout strain by mass spectroscopy, suggested that the substrate for CYP172A1 might be L-arabinose which is converted into arabinonic acid [278]. Sugar transformations are not common in P450s, one major reason being that the P450 active site is usually highly hydrophobic in nature, and so substrates are more likely to be hydrophobic molecules such as fatty acids and other lipids.
To investigate L-arabinose as a substrate, a concentrated stock solution of the sugar (250 mM) was titrated into a cuvette containing CYP172A1 (3 µM) until its concentration reached 5 mM (1000 x more than the protein concentration). No optical shift was observed (data not shown), which casts doubt on L-arabinose being a natural substrate. Corcionivoschi did not produce any UV-Visible spectroscopic results to support his mass spectrometry data, and so it is likely that the limited data presented were over interpreted. To try and identify any type I ligands for CYP172A1, a range of different potential substrate compounds were titrated against the protein (Table 5.3). None of these compounds produced any shifts, LS or HS, and so had no detectable interaction with the heme iron or the 6th (water) ligand. A range of different amino acids were among the compounds chosen, as C. jejuni is known to catabolise amino acids, including serine, as alternative carbon sources [279]. A common substrate type for P450s is fatty acids; BM3 (CYP102A1), eukaryotic CYP4 enzymes and other members of the CYP102 family are all capable of the binding and subsequent hydroxylation of fatty acids [280]. Straight chain and branched chain fatty acids were titrated against CYP172A1 but no spectral shifts were observed. Amino acid sequence alignments identified CYP201A2 from Rhodopseudomonas palustris TIE-2, the most closely related P450 with a defined substrate, with an amino sequence identity of ~30 % to CYP172A1. The unusual substrate for CYP201A2 is tributyl phosphate (TBP), but this also failed to produce a HS shift in CYP172A1 even at 1 mM concentration. A range of other structurally diverse compounds were also titrated against CYP172A1, including indene, styrene and octanoyl Co-A. None of the broad panel of compounds tested produced a substrate-like (type I) spectral shift.
Table 5.3. List of ligands tested for HS formation in CYP172A1. Compounds were mixed with CYP172A1 (3-5 µM) in a cuvette and changes in P450 spectrum measured by UV-Visible spectroscopy to probe for any changes in the heme iron spin state. Me, methyl. *Final concentration used.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Concentration*</th>
<th>Ligand</th>
<th>Concentration*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terpenes</td>
<td></td>
<td>L-Amino acids</td>
<td></td>
</tr>
<tr>
<td>Terpineol</td>
<td>140 µM</td>
<td>Arginine</td>
<td>16 mM</td>
</tr>
<tr>
<td>Linalool</td>
<td>5 mM</td>
<td>Histidine</td>
<td>10 mM</td>
</tr>
<tr>
<td>Steroids/ols</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cholesterol</td>
<td>200 µM</td>
<td>Isoleucine</td>
<td>14 mM</td>
</tr>
<tr>
<td>5-Cholest-3-one</td>
<td>150 µM</td>
<td>Glycine</td>
<td>20 mM</td>
</tr>
<tr>
<td>Estriol</td>
<td>1.2 mM</td>
<td>Serine</td>
<td>12 mM</td>
</tr>
<tr>
<td>Testosterone</td>
<td>280 µM</td>
<td>Valine</td>
<td>16 mM</td>
</tr>
<tr>
<td>Steroids/ols</td>
<td></td>
<td>Lysine</td>
<td>10 mM</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>Fatty acid</td>
<td></td>
</tr>
<tr>
<td>Octanoyl Co-A</td>
<td>2 mM</td>
<td>Lauric</td>
<td>2.5 mM</td>
</tr>
<tr>
<td>Tri butyl phenol</td>
<td>1 mM</td>
<td>Caprylic</td>
<td>1 mM</td>
</tr>
<tr>
<td>Trans-3-octene</td>
<td>2.3 mM</td>
<td>Arachidonic</td>
<td>2.0 mM</td>
</tr>
<tr>
<td>1,2 Octene oxide</td>
<td>1 mM</td>
<td>Palmitic</td>
<td>1 mM</td>
</tr>
<tr>
<td>Indene</td>
<td>12 mM</td>
<td>13-Me Myristic</td>
<td>1 mM</td>
</tr>
<tr>
<td>3-chlorostyrene</td>
<td>4.5 mM</td>
<td>15-Me Palmitic</td>
<td>1 mM</td>
</tr>
<tr>
<td>Styrene</td>
<td>1.1 mM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-arabinose</td>
<td>5 mM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.4 Compound screening to identify new ligands for CYP172A1

Initial attempts to identify a substrate, using various chemicals that have been identified as ligands for other characterised P450s, did not lead to a positive identification of any substrate-like molecules for CYP172A1. The next step taken was to screen using a much larger library of compounds. Screening was performed by the Screening Unit and Department of Medicinal Chemistry, Leibniz Institute for Molecular Pharmacology (FMP), Berlin. The library contains in excess of 20,000 synthetic organic compounds to test against CYP172A1. The screening process has already been successfully used to search for ligands for other proteins, including CYP130 from *M. tuberculosis* [281] and CYP51 from *Trypanosoma cruzi* [282]. UV-Visible spectra were analysed for changes in absorbance upon compound binding to the protein, for either type I or type II binding. A type II control, 1-PIM was used to provide a typical type II spectrum by which ‘hits’
could be identified, and compounds binding in a substrate-like manner were selected by probing for Soret spectral changes to shorter wavelengths.

5.4.1 Type I hits for CYP172A1

From the thousands of compounds screened, spectra from the 170 top ‘hits’ (based on the development of spectral troughs at ~425 nm and peaks at ~390 nm) were analysed to try and reduce the list of compounds for further investigation (i.e. rule out false positives and weak binders). From examination of spectral shifts induced, it was found that some of the potential hits gave rise to absorption difference spectra with the appropriate shape (i.e. a trough at ~425 nm and peak at ~390 nm) (Figure 5.10). Some difference spectra contained only a decrease in the Soret peak region. Another group of unsuitable compounds produced spectra with high background absorption values. In these cases, it was considered likely that the compounds caused increases to the turbidity in the protein solution, thus increasing the overall absorbance readings across the spectral region analysed.

5.4.2 Type II hits for CYP172A1

As well as looking for substrate-like molecules, the large library screen also enabled the identification of multiple novel inhibitor molecules, based on type II absorbance shifts of the CYP172A11 heme.

Despite a large number of type I compounds being identified, analysis of the spectral data produced a rather smaller number (21) of type II inhibitor-like molecules. These were identified based on an optical shift characterized by a trough at ~412 nm and a peak at ~435 nm in the difference spectrum.
Figure 5.10. Difference spectra generated by selected library compound additions to CYP172A1. A) The difference spectrum generated by compound 204073. A trough at ~425 nm and a peak at ~390 nm gives a strong indication of type I binding. B) A V-shaped curve from compound 201265, perhaps resulting from heme loss or inhibitory binding to the heme that decreased its Soret coefficient. C) A noisy background and small type I-like shift from compound 214211 indicated a possible weak binding substrate, but not as good a candidate as 204073.
5.4.3. Analysis of the interaction of novel type I and II (ChemDiv) compounds with CYP172A1

Compounds identified as those which produced strong UV-visible type I or type II shift were ranked in order of the magnitude of the spectral shift induced (i.e. the absorbance difference between the peak and trough in the Soret difference spectra). The top compounds were then, where possible, obtained from the Chemdiv compound library. In total nine compounds, eight potential substrates and one inhibitor were purchased (Appendix Figure S12).

In order to confirm the results from the compound screening experiments and to obtain dissociation constants for the binding of these molecules to CYP172A1, interactions of these hits with the P450 were followed using UV-Visible spectroscopy. Due to the high intensity colour of some of the compounds, dual beam UV-Visible spectroscopy was used to eliminate any absorbance contributions from the compounds themselves.

Binding was only observed with two of the eight substrates, 204073 and 213071. Among the other six substrates, turbidity in the UV-Visible spectra indicated that a lack of aqueous solubility was an issue for 214665 and 203650. Some of these compounds caused strong absorbance change around 390 nm, but failed to produce corresponding reciprocal changes of absorbance in other parts of the heme spectrum. Optical titrations with compound 204073 produced a typical type I shift with an almost complete conversion of the starting low spin spectrum to a high spin form (Figure 5.11), with the Soret peak at 419 nm in the absorption spectrum being replaced by one at 390 nm upon addition of 204073 to saturating levels. A shift in the low intensity heme band at 690 nm was another strong indicator that 204073 binding shifted the heme iron into a HS state. This absorption feature is the Cys-thiolate to ferric heme iron charge transfer (CT) band.

The changes in Soret absorbance induced (specifically the overall absorption difference between the peak and trough wavelengths in the difference spectra, using the same wavelength pair in each case) were plotted against their corresponding concentration of 204073, and the dissociation constant \(K_d\) calculated to be 740 ± 50 nM by fitting the data to equation 2. A small sample of 204074 was made available from the compound screening facility and it too was titrated against the CYP172A1 enzyme (Figure 5.12).
Figure 5.11. Optical titration of CYP172A1 with 204073. A) Selected UV-Vis absorption spectra of CYP172A1 (8 µM) titrated with 204073. Additions (0.25 µl) of a 204073 stock (1 mM) were added to a final concentration of 20 µM. The resting Soret peak was at 419 nm and decreased in intensity and shifted to 390 nm with addition of the 204073. Insert, shows difference spectra which were used to determine the maximum absorption changes upon the ligand binding. B) Binding curve from the titration, calculated by subtracting absorption values at the trough (420 nm) from those at the peak (388 nm) in each difference spectrum in order to obtain total absorbance change values at each [204073]. These difference values were then plotted against [204073] and the data were fitted using equation 2 to generate at $K_d$ of 0.74 ± 0.05 µM.
Figure 5.12. Optical titration of CYP172A1 with 204074. A) Selected UV-Vis absorption spectra of CYP172A1 (3.0 µM) titrated with 204074. Titrations were done with a 1 mM stock, as described in Fig. 5.11. The Soret peak shifted from 418 to 390 nm with addition of 204074. Insert shows difference spectra used to determine maximum absorption changes on ligand binding. B) Binding curve from the titration. Values were calculated by subtracting A_{320} (trough) from A_{388} (peak) for each [204074], and data plotted against [204074] and fitted using equation 2 to generate a $K_d$ of 0.75 ± 0.09 µM.
Compound 204074 was not able to produce as strong a HS shift as 204073, although the heme iron was converted to ~70% HS (with ~30% LS) at ligand saturation. Like 204073, 204074 binding resulted in the enhancement of the CT spectral feature at ~650 nm. From the binding curve generated from the difference spectra, a $K_d$ of 750 ± 90 nM was calculated. Compounds 204073 and 204074 thus bound to CYP172A1 with very similar affinities, which is due to the fact that they have very similar chemical structures (Figure 5.13). The only difference between the molecules is the presence of two methyl groups on adjacent carbons of one of the phenyl groups in 204074, while 204073 instead has an ethyl group on one of these carbons and a hydrogen on the other (blue circle). The difference in the extent of spin-state shift observed between these two compounds may reflect subtle differences in their binding mode, caused by the changes in the substituent groups. However, these changes have only a small effect on their overall affinity to CYP172A1.

Figure 5.13. Chemical structures of compounds 204073 and 204074. The chemical structures of these two molecules are very similar, with the only difference being on the N-linked phenyl ring. Compound 204073 contains an ethyl group attached to one of the phenyl carbons, but 204073 contains a methyl group at this position, and on the adjacent carbon (indicated with blue circles).
Compound 213071 was the third compound to give a substrate like shift upon binding to CYP172A1 (Figure 5.14). Binding of 213071 converted the CYP172A1 protein to HS by around 50 % (based on 390 nm peak formation). The difference absorption spectra were used to identify the position of the maximum (388 nm) and minimum (421 nm) absorption wavelengths throughout the titration. The total absorption differences between the peak and trough values in the spectra collected at the different concentrations of 213071 were then plotted against the corresponding [213071]. Data were fitted using equation 2 and a $K_d$ of 2.09 ± 0.1 µM was determined. The structure of 213071 is quite similar to 204073/4 and so it was not surprising that a similar dissociation constant was calculated (Figure 5.15).

**Figure 5.15. Chemical structures of 204073 and 213071.** The two compounds have similar size, but 213071 (right) has a thiophene instead of a toluene (methylbenzene group) on 204073 (left). A methyl group in 213071 replaces the ethyl substituent on the N-linked phenyl group in 204073, and is attached via an imidazole rather than a 2-methylpyridine moiety.
Figure 5.14. Optical titration of CYP172A1 with 213071. A) Selected UV-Vis absorption spectra of CYP172A1 (3.1 µM) titrated with 213071. Titrations were done with a 1 mM stock, as described in Figure 5.11. The Soret peak shifted from 418 to 390 nm with addition of 213071. Insert shows difference spectra used to determine maximum absorption changes on ligand binding. B) Binding curve from the titration. Values were calculated by subtracting A_{421} (trough) from A_{388} (peak) for each [231071], and data plotted against [213071] and fitted using equation 2 to generate a $K_d$ of 2.19 ± 0.10 µM.
The only available inhibitor molecule (from the screen) available from ChemDiv was 204591. During the titration with 204591 there was an increase in the absorbance intensity in the δ band region (360 nm), suggesting that the enzyme retained its LS state (Figure 5.16). The main LS Soret band intensity (419 nm) dropped sharply with each addition of 204591, but retained its approximate wavelength maximum. A shoulder developed at 445 nm during the titration. Difference spectra were generated as before. The maxima and minima from the difference spectra were identified and the overall absorbance change at the same wavelength pair (between 446 nm and 421 nm) was plotted against [204591]. Equation 2 was used to determine a dissociation constant of 1.22 ± 0.07 µM. The type of shift observed was not typical of that seen for nitrogen based ligands (e.g. azoles), perhaps suggesting that the 204591 interaction with the heme did not occur via a nitrogen atom, but possibly instead, via a sulphur atom. Such bis-thiol(ate) coordination of the heme iron is not common, but is suggested to occur in the DGCR8 protein involved in microRNA biosynthesis. This enzyme has a resting ferric heme Soret maximum at 450 nm [283].

The four ChemDiv compounds which bound to CYP172A1 were all found to have $K_d$ values in the low micro-molar range or tighter (Table 5.4). These are much tighter $K_d$ values than for the other ligands that bound to CYP172A1 (Table 5.2). The bulkier structures of the ChemDiv compounds may enable more favourable hydrophobic, hydrogen bonding and other interactions within the active site to enable their tighter binding. The subtle structural differences with the three type I compounds appeared to illustrate the importance of matching the topology of the active site to the chemical structures. Small variation in their structure e.g. by a change from an ethyl group to two methyl groups altered the degree of the high spin shift from ~95 % to ~70 % in the case of 204073 and 204074, respectively.
Chapter 5. Ligand binding to CYP172A1

<table>
<thead>
<tr>
<th>Compound</th>
<th>Soret peak shift (nm)</th>
<th>HS:LS</th>
<th>$K_d$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>204073</td>
<td>418 - 390</td>
<td>95:5</td>
<td>0.74 ± 0.05</td>
</tr>
<tr>
<td>204074</td>
<td>418 - 390</td>
<td>70:30</td>
<td>0.75 ± 0.09</td>
</tr>
<tr>
<td>213071</td>
<td>418 - 390</td>
<td>50:50</td>
<td>2.09 ± 0.1</td>
</tr>
<tr>
<td>204591</td>
<td>418 - 445</td>
<td>0:100</td>
<td>1.22 ± 0.07</td>
</tr>
</tbody>
</table>

Table 5.4. ChemDiv substrate and inhibitor compounds and their affinity for CYP172A1. 204073, 204074 and 213071 all exhibited a type I binding shift with CYP172A1, with the Soret peak shifting towards 390 nm and development of HS heme iron. 204591 produced a type II spectrum and a new LS spectrum with a Soret shift to 445 nm. The CYP172A1 dissociation constant for each molecule is provided.
Figure 5.16. Optical titration of CYP172A1 with 204591. A) Selected UV-Vis absorption spectra of CYP172A1 (3.5 µM) titrated with 204591. Titrations were done with a 1 mM stock, as described in Figure 5.11. The ferric Soret peak at 419 nm decreased in intensity upon ligand binding and a shoulder appeared at 444 nm with addition of the 204591. Insert shows difference spectra used to determine maximum absorption changes on ligand binding. B) Binding curve from the titration. Values were calculated by subtracting $A_{418}$ (trough) from $A_{447}$ (peak) for each [204591], and data plotted against [204591] and fitted using equation 2 to generate a $K_d$ of 1.22 ± 0.07 µM.
5.5. EPR of ligand-bound forms of CYP172A1

EPR analysis of CYP172A1 in different ligated-bound forms was done to further analyse the types of inhibitor- or substrate-like interactions that occurred. The EPR signal of the native protein (Figure 5.17, black line) indicated a single species (as discussed in Section 4.7). On addition of econazole there was the generation of a new (major) species with g values at 2.44 ($g_z$), 2.24 ($g_y$) and 1.89 ($g_x$) and a minor species with g values at 2.41, 2.24 and 1.91 (red). The former g value set was consistent with nitrogen coordination of the heme iron, with the latter set likely being from a proportion of the enzyme that retains the water 6th distal ligand. The g values might actually indicate that econazole does not displace the heme iron distal water. In some cases, an imidazole nitrogen directly ligating to a P450 heme iron could produce a $g_z$ value nearer to 2.5 [239]. Thus, it is possible that econazole binds to the heme via an internal water molecule retained as the heme 6th ligand. In the case of CYP51 in M. avium the $g_z$ values for econazole and ketoconazole complexes were both at 2.46 [284]. However, the crystal structure of CYP121, from M. tuberculosis, complexed with fluconazole showed azole coordination via a bridging water molecule [269].

Binding of 1-PIM (red) and 4-PIM (black) also produced a major and a minor LS species with corresponding g values. Interestingly, the 1-PIM creates a major LS species with g values of 2.44, 2.24 and 1.89 (and minor species of 2.41, 2.24 and 1.91), whereas the 4-PIM has major g values of 2.41, 2.24 and 1.91 (and minor species of 2.44, 2.24 and 1.89). The 2.41, 2.24 and 1.91 set is essentially the same as the resting form (water ligated), and thus the second species is each case is that of the PIM ligating to the heme iron (either directly or via the 6th water ligand). It is interesting that 1-PIM and 4-PIM displayed differing EPR spectra, despite their structural similarity and both having $K_d$ values of ~40 µM. The differences in EPR are potentially related to the subtly different binding modes and the low temperatures used for collection of the EPR spectra. The 1-PIM EPR spectrum reveals a higher proportion of the azole coordinated directly to the heme iron (or via the 6th distal water ligand), whilst in the 4-PIM this is only a minor species. The coordination of these compounds to the heme iron (direct or water mediated) could be determined by X-ray diffraction studies on inhibitor-bound CYP172A1 crystals.
EPR data were also collected for CYP172A1 bound to compounds acquired from ChemDiv that displayed inhibitor (204591) or substrate (204073 and 213071) type binding (Figure 5.18). The 204073 compound had a HS set of g values of 8.00, 3.45 and 1.65 [285]. There was also a LS species formed with g values of 2.37, 2.23 and 1.97. The HS set is consistent with the HS-like optical shift observed at ambient temperature. The g values obtained are distinct from those obtained for the ligand-free CYP172A1, and are likely a consequence of the 204073 being bound in the P450 active site cavity in close proximity of the heme iron. The LS species form may result from the influence of the cryogenic temperatures employed in order to obtain in the EPR spectrum.

Compound 213071 also produced a spectrum consistent with a HS species with g$z$ and g$y$ values of 8.03 and 3.39, respectively. The signal for the g$x$ value (1.61) was very weak and only just appears out of the background level. A low spin species was also present with g values of 2.41, 2.25 and 1.92. The binding of 213071 does not produce a full HS shift as observed by UV-Visible spectroscopy (Figure 5.14), and this LS set is similar to the observed g values for the ligand-free CYP172A1. Potentially the binding mode of 213071 (at least under EPR temperatures) is such that a water molecule is able to ligate again as the 6th ligand to the heme iron.

The EPR spectrum for the inhibitor molecule 204591 revealed a major species with g values of 2.41, 2.25 and 1.92. The g values are similar to those for the native protein, which may indicate that at 10 K 204591 does not coordinate the heme iron, and that a water ligand is retained. A minor species with g values of 2.44, 2.24 and 1.89 was present in the trace. This would be consistent with the water being replaced by another ligand. It was not possible from these values to determine if the coordination was from a nitrogen present on the molecule or, as the UV-Visible absorption data might suggest, from a sulphur atom (Figure 5.19).
Figure 5.17. EPR spectra of ligand-bound and ligand-free forms of CYP172A1 A) EPR of CYP172A1 ligand-free (black) displays a single low spin species with g values of 2.41, 2.24 and 1.91. On binding, econazole (red) two species exist. The minor form has the same g values as the ligand-free state, but a new major species is seen with a set of g values at 2.44, 2.24 and 1.89. B) EPR spectra for CYP172A1 bound with 1-PIM (red) and 4-PIM (black). Both complexes exist in low spin states but in differing mixtures of forms with the same g values as those reported for the ligand-free and econazole-bound states reported previously.
Chapter 5. Ligand binding to CYP172A1

Figure 5.18. EPR spectra of ligand-bound forms of CYP172A1. EPR spectra were recorded for CTP172A1 bound to three ligands identified from the compound screening studies. The two substrate-like compounds, 204073 (black spectrum) and 213071 (red spectrum) and the inhibitor 204591 (green spectrum) all gave spectral shifts consistent with their binding to CYP172A1. Both the 204073 and 213071 compounds show shifts to HS states with g values around 8.0, 3.4 and 1.65 (the lowest g value is not marked in the case of 213071, but is present as a low intensity feature). The 204591-bound CYP172A1 has two sets of g values at 2.41, 2.25 and 1.92 major and 2.44, 2.25, and 1.89 (minor) indicating two low spin species. The former is likely a state with the water retained as the 6th ligand, while the latter should originate from heme iron ligation by the S- or N-atom from the 204591 compound.

Figure 5.19. Chemical structure of the 204591 compound. The structure of the polycyclic 2040591 structure contains both sulphur and nitrogen atoms that could coordinate to the heme iron.
5.6. Redox potentiometry of CYP172A1

In order to investigate the heme Fe(III)/Fe(II) midpoint potential of CYP172A1, redox potentiometry (spectroelectrochemistry) was carried out under anaerobic conditions, as the presence of oxygen would lead to reoxidation of the heme, preventing full heme reduction. Determining the heme iron redox potential allows analysis of the ability of the heme iron to accept electrons required for catalysis. Potentiometry was carried out using an adaption of the method devised by Dutton [187]. Small volumes of a concentrated solution of sodium dithionite were titrated into the CYP172A1 solution and the redox state of the P450 heme allowed to reach equilibrium before, simultaneously, recording the applied potential and the P450 spectrum. The reduction of the heme iron was accompanied by absorbance changes in the heme (Figure 5.20). The Soret peak at 418 nm, from the LS heme species, decreased in intensity and a new, less intense peak at 408 nm emerged, caused by the formation of the ferrous heme. The α and β bands merged spectrally as the redox potential decreased below -400 mV and heme reduction neared completion.

The change in the heme absorbance at 418 nm was plotted against the applied redox potential and the data were fitted using the Nernst equation to produce a midpoint potential for the heme iron Fe(III)/Fe(II) couple. The midpoint potential value of -418 ± 4 mV (with reference to the standard electrode SHE) is similar to those of various other P450s in their substrate-free form, including CYP102A1 (-427 mV) [286].

The heme iron redox potential of CYP172A1 with 204073 (a substrate-like molecule identified from compound screening, Section 5.4) bound was also determined to establish whether a positive shift in the heme iron midpoint potential was observed, as has been reported for several other P450s, including, CYP102A1, CYP51B1 (a M. tuberculosis sterol demethylase) and CYP101A1 (the P. putida camphor hydroxylase) [241]. With the 204073 molecule bound, the ferric enzyme was converted to an extensively HS state (with Soret maximum at ~390 nm). The reduction of 204073-bound CYP172A1 with sodium dithionite results in a heme Soret peak shift to 410 nm (similar to that for the substrate-free form) (Figure 5.21).
Figure 5.20. Redox potentiometry of substrate-free CYP172A1. A) Selected UV-Vis absorption spectra of CYP172A1 (9 µM) collected during redox titration with sodium dithionite. The completely oxidized form (black) is almost completely reduced (magenta) upon the final addition of sodium dithionite. The Soret band shifts from 418 nm to 408 nm upon heme reduction. The α and β bands (at 568 and 534 nm, respectively, for the oxidized P450) merged into a single species at 549 nm as the iron became reduced. At 465 nm there was also an increase in absorption as the heme iron was reduced and the Soret feature broadened. The changes in absorption at 418 and 406 nm thus reflected a ferric to ferrous heme iron reduction. B) A plot of the $A_{418}$ against the applied potential. The Nernst equation was used to fit the data, and produced a midpoint potential for the heme iron $E^{'''} = -418 \pm 4$ mV.
For the 204073-bound form of CYP172A1, the change in heme absorbance at 390 nm was plotted against the applied redox potential and the data fitted using the Nernst equation. This produced a midpoint potential of $-379 \pm 1$ mV vs. SHE. This positive shift of $\sim 40$ mV was lower than those observed for certain other P450s, which are often greater than 100 mV [241, 286]. The extent of the change may be underestimated slightly due to a slightly inaccurate estimation of the substrate free-free potential, where reduction of CYP172A1 may not have gone to completion. In addition, the 204073 is not a natural substrate for the enzyme, and so it is quite possible that the real substrate might further increase the shift to midpoint potential, possibly above that of the likely electron donor cofactor NAD(P)H (-320 mV) [287].

Upon reoxidation of the reduced forms of CYP172A1 using potassium ferricyanide, there was no apparent hysteresis (i.e. spectra recorded at the same potential as those collected during the reductive reaction were highly similar), suggesting that the process of CYP172A1 heme reduction/oxidation is freely reversible. In the case of other P450s, CYP142 from M. tuberculosis displayed hysteretic behaviour suggested to originate from structural rearrangements of the protein in its conversion between ferric and ferrous states [240].
Figure 5.21. Redox potentiometry of 204073-bound CYP172A1. A) A UV-Vis absorption spectra of CYP172A1 (9 µM) bound with 204073 (100 µM) during redox titration with sodium dithionite. The spectrum in grey is that for the oxidized substrate-free CYP172A1, and upon binding of 204073 (red) the Soret maximum shifts to 390 nm. On heme reduction the Soret band shifts to 410 nm (magneta). In the Q-band region, the distinct α and β bands in the ferric state are converted into a single spectral species at 549 nm as the iron was reduced. Increases in absorption are noted in the region from 400-500 nm, with a substantial increase at 465 nm as the heme iron was reduced. The Soret changes in absorption at 390 and 410 nm reflect the ferric to ferrous heme reduction for the substrate-bound P450. B) A plot of the Soret A390 against the applied potential for spectra collected during the redox titration. The Nernst equation was used to fit these data, generating a midpoint potential for the heme iron of $E'' = -379 \pm 5$ mV.
5.7. CO trapping of ferrous CYP172A1 in the presence of redox partners

P450s require redox partners to facilitate the transfer of electrons in order for them to carry out oxidative catalysis. However, the natural redox partners are not currently known for CYP172A1. Fortunately, it is well known that non-natural redox partners (i.e. NAD(P)H-dependent flavoprotein reductases and ferredoxin/flavodoxins) can act as electron donors to heterologous P450s [240, 288]. E. coli flavodoxin (FldA) and flavodoxin reductase (FldR), and spinach ferredoxin (spFdx) class I redox systems were used to facilitate the NADPH-dependent reduction of CYP172A1, with binding of CO to the heme protein then trapping the P450 in a stable Fe(II)CO complex. By following the kinetics of the process in the heme Soret region, kinetics of the electron transfer from NADPH through the heme iron were measured, since the kinetics of CO binding to the ferrous heme iron is certain to be faster that the electron transfer reactions through the redox partners. The reductant sodium dithionite was also used to transfer electrons to CYP172A1 in a redox partner-free system for comparison. All experiments were conducted in a glove box to prevent reoxidation of CYP172A1 prior to CO binding, and the 204073 substrate molecule was also added in a parallel series of experiments in order to examine the kinetics of P450 reduction in the presence of a substrate-like molecule.

In the presence of sodium dithionite the Fe(II)CO complex was rapidly formed (Figure 5.22). The resting (ferric) protein’s Soret peak at 418 nm (blue solid line) rapidly diminished in intensity and almost complete conversion to the P450 form was observed (red solid line, peak at 448 nm). There was a small amount of P420 formed, likely indicating a proportion of the heme iron being cysteine-thiol coordinated. The distinct α and β bands seen in the ferric state merged into a broad peak (~553 nm) upon heme reduction. By plotting the change in absorbance, at 448 nm, against time and fitting the data using an exponential function, a rate constant for P450 formation was calculated to be 0.297 ± 0.026 min⁻¹. The complex was very stable and the oxidative reconversion to resting, ferric state, was not complete after 20 hours (Figure 5.22). The progress curve for the collapse of the P450 form was linear, and the rate constant for the process was estimated to be approximately 1x10⁻⁴ min⁻¹, which is 3000-fold slower than the rate of the P450 formation.
Figure 5.22. Trapping of the Fe(II)CO complex of CYP172A1 using sodium dithionite. A) UV-Vis spectra of CO binding to dithionite-reduced CYP172A1 (2.7 µM). The ferric Soret peak at 418 nm was lost as the new peak emerged at 448 nm, consistent with the formation of a Fe(II)CO complex. The α and β bands merge to form a new broad peak (553 nm). Insert, rate of Fe(II)CO complex formation. The change in the absorbance (448 nm) was plotted against time. A single exponential function was used to fit the data to generate a rate constant of 0.297 ± 0.026 min⁻¹. B) UV-Vis spectra showing the process of conversion of the P450 Fe(II)CO complex of reduced CYP172A1 (2.7 µM) back to the ferric state. The Soret peak at 418 nm increases in intensity over time. The α and β bands begin to emerge from the broad peak (553 nm), consistent with the reoxidation of the Fe(II)CO complex back to the resting ferric state. Insert, rate of Fe(II)CO complex collapse. The change in the absorbance (448 nm) was plotted against the time. A linear fitting process was used to estimate the rate constant for the P450 complex collapse to be 1.04 ± 0.01 x10⁻⁴ min⁻¹.
CO trapping provides an elegant way to observe electron flow from redox partners to the heme iron, as the CO will only bind to reduced (ferrous) heme iron. As the natural redox partners for CYP172A1 were unknown, CO trapping was used to investigate if CYP172A1 could interact with non physiological redox partners. A reaction mixture of CYP172A1:FldR:FldA (1:2:5 stoichiometry) was designed (with the expectation that an excess of redox partners should enhance electron transport to the heme iron), as electron transfer kinetics were expected to be a rate limiting step in the reaction, while CO binding to ferrous heme iron should be rapid. The reaction buffer was saturated with CO, and the reaction started with the addition of excess NADPH (100 µM). The effect of adding a 204073 (100 µM) on the rate of P450 formation was also investigated under the same conditions.

The two *E. coli* redox partners were able to transfer electrons to CYP172A1 (Figure’s 5.23 and 5.24), and this produced a red shift of the Soret peak to ~448 nm, consistent with the Fe(II)CO complex formation. Even in the absence of substrate, the CO was able to bind to the reduced heme to form the Fe(II)CO complex. An accurate indication of the initial rate of Fe(II)CO formation was not available, as the absorption of the flavins in FLDR and FldA changes over the first period of the reaction until an equilibrium state is reached (with a substantial amount of flavin semiquione formed, likely mainly from the flavodoxin, and that obscures the heme absorption in the Q-band region). Following this, the distinguishable conversion to P450 species at 448 nm was not complete, which may reflect the non physiological conditions of the experiment and/or the slow electron transfer kinetics, and/or the formation of a proportion of the Fe(II)CO P420 complex. By following the change of the absorbance at 448 nm against time it was possible to determine the rate of the P450 formation, and also to estimate the rate constant for P450 complex collapse. In the substrate-free state, the rate constant for Fe(II)CO complex formation was $4.12 \pm 0.050 \times 10^{-3}$ min$^{-1}$, which was slower than the rate at which the complex formed in dithionite reduced CYP172A1. Using *E. coli* redox partners, it took ~500 minutes for the Fe(II)CO complex to be maximally formed. The complex was stable (i.e. reached equilibrium) for around 100 minutes before notable collapse of the complex to the ferric form and/or proportions of the P420 state began. The rate constant of Fe(II)CO complex collapse to the P420 form was $1 \times 10^{-4}$ min$^{-1}$,
rather slower than the rate constant for formation of the P450 complex. The reaction progress was essentially linear. In experiments done with 204073 (100 µM) bound (so that the P450 had a substantial amount of HS heme iron) the comparable rate constant for P450 formation was $1.214 \pm 0.007 \times 10^{-2}$ min$^{-1}$ which was three times quicker than for the substrate-free form. This might be associated with the positive change in the heme iron redox potential of the substrate bound enzyme (as shown in Section 5.6) and also possibly due in part to structural reorganization of the P450 (conducive to partner interaction) that may occur in the substrate-bound state. The rate of conversion to reoxidized (ferric) P450 with CO dissociation/P420 state was also slower than for the substrate free enzyme ($3.93 \pm 0.02 \times 10^{-5}$ min$^{-1}$) suggesting that the P450 complex was more stable in the 204073-bound form than in the substrate-free state.
Figure 5.23. Formation and collapse of Fe(II)CO complex by CYP172A1 using the *E. coli* Fldr/Flda redox partners. A) UV-Vis spectra showing CO binding to CYP172A1 (2.7 µM) following heme iron reduction by the redox partner system. The ferric Soret peak at 418 nm diminished and a new Soret feature emerged at 448 nm (formation of Fe(II)CO complex). The oxidised Fldr/Flda spectra have typical flavin features across the region from 300-500 nm. These decrease in intensity during early stages of the reaction as the flavins reach an approximate equilibrium state of partial reduction. The spectrum for substrate-free (black dashed line) P450, and for CYP172A1 mixed with redox partners and NADPH immediately after the start of the experiment (red line) and at the end point of the experiment (500 minutes, blue line) are shown. Insert, the change in the absorbance (448 nm) was plotted against the time. A single exponential fit was used to generate a rate constant of $4.12 \pm 0.05 \times 10^{-3}$ min$^{-1}$ for Fe(II)CO complex formation. B) UV-Vis spectra reporting Fe(II)CO complex collapse from the P450 Fe(II)CO state of CYP172A1 (2.7 µM). The Soret peak at 418 nm increased in intensity over time, probably consistent with the reoxidized ferric CYP172A1 being the main product formed. The α and β bands (at positions typical of the resting, ferric state) emerged from the broad peak (553 nm) of the reduced enzyme. Insert, the change in the absorbance (at 448 nm) was plotted against time. A linear fit was used to estimate the rate of Fe(II)CO complex collapse to be $\sim 1.14 \pm 0.01 \times 10^{-4}$ min$^{-1}$.
Figure 5.24. Formation and collapse of Fe(II)CO complex of CYP172A1 in presence of 204073 substrate using the E. coli FldR/FldA redox partners. A) UV-Vis spectra showing the progressive CO binding to reduced CYP172A1 (2.7 µM). The Soret peak at 390 nm was lost and a new peak emerged at 448 nm (formation of Fe(II)CO P450 complex). The reaction proceeded as described in Figure 5.23. Insert, rate of Fe(II)CO P450 complex formation. The change in the absorbance (at 448 nm) was plotted against time. A single exponential fit was used to generate a rate constant of $1.214 \pm 0.007 \times 10^{-2} \text{ min}^{-1}$. B) UV-Visible spectra following CO release from the reduced/CO bound CYP172A1 P450 complex (2.7 µM) and P450 reoxidation. The Soret peak at 418 nm increased in intensity over time as the P450 complex converted back towards the ferric state. Insert, rate of Fe(II)CO complex collapse at 448 nm. The change in $A_{448}$ was plotted versus the time. A linear fit was used to estimate the rate of collapse of the P450 complex to be $3.93 \pm 0.02 \times 10^{-5} \text{ min}^{-1}$.
Different types of flavodoxin and ferredoxin proteins (and flavodoxin/ferredoxin reductases) support the function of a large number of bacterial (and mitochondrial) P450s. To investigate whether there was a preference of CYP172A1 to accept electrons from spinach ferredoxin (Fdx), the redox component system was made up in the ratios as described previously, but using spFdx instead of the flavodoxin. The rate of Fe(II)CO complex formation was compared in the presence and absence of 204073.

The FldR:spFdx redox system was able to pass electrons to the heme, which allowed the binding of CO to the ferrous heme iron, both with and without 204073 present (Figure 5.25). As the Fdx does not have a strong UV absorbance between 300 to 500 nm, the P450 formation was observed much more clearly than in the situation where FldA was used as the P450 redox partner. In the presence or absence of the 204073, almost complete conversion to P450 form was observed, with a small amount of P420 also appearing.

The change in absorbance at 448 nm under these conditions was plotted against time to generate a plot, with the data fitted using a single exponential function. The time required for the reaction to go to completion (i.e. maximal P450 formation) was ~100 minutes with 204073-bound enzyme (in a predominantly HS state) and around 200 minutes for the substrate-free (LS state). The rate constant was determined to be $2.87 \pm 0.04 \times 10^{-2} \text{ min}^{-1}$ with 204073 bound, which was 1.5x faster than the ligand-free rate constant ($1.81 \pm 0.02 \times 10^{-2} \text{ min}^{-1}$).

Comparing all the different conditions tested, it appears that the *E. coli* FldR redox partner is a more efficient reductant of CYP172A1 with spFdx than it is with the *E. coli* FldA. Increases in the rate constants were observed for both partner systems on the addition of the 204073 substrate, consistent with its effects on the HS heme iron development and increase in the heme iron reduction potential (Table 5.5). Sodium dithionite is able to reduce the P450 heme directly, and thus the CYP172A1 P450 complex can form quickly, even in the absence of a substrate.
<table>
<thead>
<tr>
<th>Redox system</th>
<th>204073 present</th>
<th>Fe(II)CO complex $A_{\text{max}}$</th>
<th>Rate of P450 formation (min$^{-1}$)</th>
<th>Rate of P450 collapse (min$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium dithionite</td>
<td>No</td>
<td>448 nm</td>
<td>$0.297 \pm 0.026$</td>
<td>$1.04 \pm 0.01 \times 10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>448 nm</td>
<td>$1.214 \pm 0.007 \times 10^{-2}$</td>
<td>$3.93 \pm 0.02 \times 10^{-5}$</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>447 nm</td>
<td>$4.12 \pm 0.05 \times 10^{-3}$</td>
<td>$1.14 \pm 0.01 \times 10^{-4}$</td>
</tr>
<tr>
<td>FldA/FldR</td>
<td>Yes</td>
<td>448 nm</td>
<td>$2.87 \pm 0.04 \times 10^{-2}$</td>
<td>ND</td>
</tr>
<tr>
<td>spFdx/FldR</td>
<td>Yes</td>
<td>448 nm</td>
<td>$1.81 \pm 0.02 \times 10^{-2}$</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>448 nm</td>
<td>$1.14 \pm 0.01 \times 10^{-4}$</td>
<td>ND</td>
</tr>
</tbody>
</table>

Table 5.5. Comparison of different redox partners with respect to rate constants for P450 complex formation and collapse for CYP172A1. P450 complex formation was identified by the shift in the heme Soret peak towards ~ 450 nm. The effect on the rate of P450 formation and collapse in the presence of 204073 was also investigated.
Figure 5.25. Formation the of Fe(II)CO complex of CYP172A1 in presence and absence of 24073 using the *E. coli* FldR and SpFdx. A) UV-Vis spectra of CO binding to CYP172A1 (2.7 µM) bound with 204073 (100 µM). The Soret peak at 390 nm diminished and the Soret peak for the Fe(II)CO complex emerged at 448 nm. The reaction proceeded as described in Figure 5.23. The initiation of experiment (red line), and the end point of the experiment where maximal Fe(II)CO complex was maximal (125 minutes, blue line) are indicated. Insert, rate of Fe(II)CO complex formation. The change in the absorbance (at 448 nm) was plotted versus time. A single exponential function was used to fit the data, generating a rate constant of $2.87 \pm 0.04 \times 10^{-2}$ min$^{-1}$ for Fe(II)CO complex formation. B) UV-Vis spectra following trapping of reduced/CO bound complex of CYP172A1 (2.7 µM) in absence of ligand. The Soret peak at 448 nm increased in intensity over time. Insert, the change in the absorbance (at 448 nm) was plotted against time. A single exponential function was used to fit the data and to generate a rate constant of $1.81 \pm 0.02 \times 10^{-2}$ min$^{-1}$ for Fe(II)CO complex formation.
5.8. Turnover of ChemDiv compounds by CYP172A1

As several compounds from the compound screen indicated type I binding to CYP172A1 investigations into whether products could be formed from one or more of these molecules were performed. Initially, turnover studies were performed using the *E. coli* FldR/FldA and FldR/spFdx redox partners, as they are able to be able to transfer electrons to the CYP172A1 heme iron. There were no observed differences in product formation observed between the two redox systems using 204073, and so the turnover experiments were then carried out using the FldR/FldA redox system. NADPH, in a large excess (1 mM), was used to deliver the required reducing equivalents and NADP* was re-reduced by the Glucose-6-Phosphate (G6P)/G6P-dehydrogenase (G6PDH) recycling system. To initiate the reaction, 100 µM of substrate was added and the mixture stirred O/N at room temperature, or for two hours at 37 °C. The reaction was then stopped with the addition of dichloromethane (DCM), which also extracted product and substrate from the soluble enzyme/cofactor aqueous phase. The DCM fraction was removed as the solvent was evaporated to concentrate any recovered substrate/product. The material was then redissolved into a methanol/acetonitrile (MeOH:ACN) (50:50) mixture containing 0.01 % (v/v) formic acid.

The samples were then analysed by mass spectrometry to identify the substrate ion and to look for any product ions. Any samples containing potential products were then analysed by LC-MS (method described in Section 2.45). This allowed for the separation of substrate and potential product peaks whose identities could be ascertained by mass spectrometry.

From the initial results it became apparent that the choice of buffer had to be carefully planned, as the most stabilising buffer (100 mM KP₆, with 250 mM KCl, pH 8.0) produced MS traces with substrate adducts of + 38, i.e. the *M*ᵣ of a potassium ion. The buffer composition was thus changed to 100 mM HEPES with 100 mM NaCl at pH 7.5. In this buffer no sodium or other adducts were observed.
5.8.1. Turnover of 204073
Despite its tight binding, a product ion could not be observed for this compound (Appendix, Figure S13). It is possible that the compound fits the active site pocket and induces a HS shift, but lacks an appropriate atom or functional group positioned above the heme, such that turnover was an undetectable level, or that it binding mode was such that it did not allow access of oxygen to the heme iron when the heme becomes reduced.

5.8.2 Turnover of 213071
In view of the inability to detect a product from the 204073 substrate, further studies were done using the structurally related 213071 molecule. Initial analysis using mass spectrometry identified a peak at 351 m/z, which was +16 compared to the substrate peak (335 m/z). This peak was not present in control samples, e.g. when the protein was boiled prior to addition to the reaction mixture, indicating that the production of this potential product peak was a result of CYP172A1 enzyme addition. The 351 peak was small, typically around 5 % compared to the substrate (which was the largest identified peak and set as 100 % intensity). The sample was then analysed by LC-MS (as described in Section 2.45). Using a 5-75 % ACN gradient the material did not elute until after 12 minutes (Figure 5.26). This first peak (b) eluted at 13 minutes and was identified as the substrate as the mass spectrum contained a single ion at 335 m/z. The next peak (c) eluted at 13.75 minutes and was a mixture of four species. The first was some residual substrate which may have been carried over from peak b. The second species had a m/z of 351 which would be the substrate + 16 peak observed in the previous analysis. The final two species had not been observed previously and had m/z values of 353 and 369, + 18 and + 34 respectively, compared to the substrate. The peak shapes looked very similar to those seen for the substrate.

The atom in the 213071 compound that undergoes oxidation is not known. However based upon results collected to date, the sulphur atom is the most likely candidate. The 351 m/z ion could be produced if a sulfoxide (S=O) was produced. The 353 and 369 species could be created if sulfenic acid (S-OH) and sulfinic acid (SOOH) derivatives were produced, respectively (Appendix Figure S14).
Figure 5.26. LC-MS analysis of 213071 turnover by CYP172A1. A) LC trace following the extraction of material following a 2 hour turnover experiment. B) Mass spectrometry analysis of the peak marked B from the LC trace. The 334.9 m/z matches that of the starting compound 213071. C) Mass spectrometry analysis of the peak marked C from the LC trace. A mixture of compounds was identified at m/z 335, 351, 352.9 and 369.
In an attempt to achieve a greater resolution of the four species from the elution profile on the LC, conditions were altered to make the gradient of increasing ACN shallower around the conditions where the compounds eluted (Section 2.45). This meant increasing the starting ACN composition to 35 % and increasing it to 55 % over the 20 minute elution. The improved resolution meant that mass spectrometry analysis of the well resolved peaks was improved.

An improved separation was achieved whereby the 351 m/z species (potentially the sulfoxide product of the substrate) was isolated from the other two compounds (arrow in Figure 5.28). There were two minor compounds at m/z 242 and 335, the latter being unmodified 213071 and the former being an unknown contaminant that forms a prominent peak at seven minutes. MS/MS of the 351 peak produced two fragments at, 244 and 228 m/z. The peak at 6.4 minutes was still a mixture of the 353 and 369 ion species and so MS/MS could not be undertaken.

5.9. *C. jejuni* sensitivity to 213071, 204074 and 204591 compounds
The compounds 204073, 204591 and 213071 were assayed to see if they had any effects of the growth of *C. jejuni*. Although they are not natural compounds, it was interesting to access if they influenced metabolism and/or had any toxic effects that might be mediated through interactions with CYP172A1. As in Chapter 3.10, the disc assay method was used. A range of concentrations of each compound (0.2 – 50 mg/ml) were used. *C. jejuni* was grown on MH agar plates in the presence of the compounds for 48 hours. After this time the plates where checked for any signs of growth inhibition.

When compared to the DMSO control, all cell types (PE1, PE2, PE4 and PE5) displayed no growth impairments. This might be due to problems with uptake of these molecules into *C. jejuni* cells, since these compounds are highly apolar and there may be issues with their solubility as well as cell penetration.
Figure 5.27. LC/MS trace of 213071 turnover reaction by CYP172A1. A) Elution profile produced from CYP172A1/213071 reaction mixture resolved on a C18 column. Three peaks are observed at t = 5.2 (unmodified substrate), 6.4 (possible sulfenic and sulfinic acid derivatives) and 6.8 (sulfoxide product) minutes. The peak of interest is marked with an arrow. B) MS profile of selected peak from LC elution shows a predominant 351 m/z ion. C) MS/MS of the 351 peak produces two peaks at 228 and 244 m/z.
5.10. Summary

A diverse range of compounds were investigated to look for type I (substrate-like) and type II (inhibitor-like) binding ligands for CYP172A1. Following the identification of several substrate-like molecules from a broad compound screen, a series of experiments involving electron transfer and thermodynamic analysis were undertaken to investigate the effects of identified ligands on the biochemical properties of the enzyme.

Azoles are well known inhibitors of P450s, and various azole scaffolds have been developed to inhibit the fungal CYP51 enzymes, which are sterol 14-α-demethylases. A range of azoles were titrated against CYP172A1 and all of those tested produced inhibitor binding characteristics when examined by UV-Visible spectroscopy, whereby the Soret peak undergoes an absorbance red shift as the heme is coordinated by an imidazole or triazole nitrogen and the low spin ferric heme is stabilised. Econazole and miconazole had the best $K_d$ values, which were both in the low nanomolar range, at 100 nM and 170 nM, respectively. Both share very similar structures and only differ by one chlorine atom. The two large hydrophobic rings presumably make favourable interactions with hydrophobic residues that are located in the core of the protein to enhance the affinity of these inhibitors for CYP172A1. Imidazole is another compound that binds as an inhibitor to P450s, albeit generally quite weakly. The observed $K_d$ (3.28 mM) was much weaker than those for the azoles drugs which highlighted the importance of the hydrophobic side groups in making strong ligand interactions with the protein. Imidazole is a small polar molecule (unlike the azole drugs) and so binds relatively weakly to the active site, as its binding is disfavoured in the apolar environment since it cannot form many favourable interactions with surrounding amino acids. The addition of a phenyl ring to the imidazole (as in the case of PIMs) increases the tightness of binding to the heme iron by ~100 fold in the case of 1- and 4-PIM. 2-PIM did not bind to the CYP172A1 heme iron, possibly due to an alternative binding mode within the active site, or to steric obstruction of the imidazole group preventing its binding to the heme iron. Hydrophobic interactions between apolar side chains in the CYP172A1 active site and the phenyl ring are a likely explanation for the enhanced affinity of the 1-PIM and 4-PIM over imidazole.
Though many type II molecules were indentified (including the azoles), the detection of type I ligands proved more difficult. With no similar P450 enzymes identified with known substrates that could be used as a starting point, it was not feasible to test with any confidence molecules that were likely substrates or related compounds. Instead a range of ligands that are known substrates for other P450s were initially tried, including linalool, a range of fatty acids and cholesterol. Due to the lipophilic nature of typical P450 substrates, these compounds suffer from poor solubility in aqueous solutions. However, binding to the apolar P450 active site, which is also lipophilic can sequester the compound from solution. Despite this theory, no binding was observed with a wide range of compounds known to be substrates for other P450s. A previous report had suggested L-arabinose as the natural substrate for CYP172A1 [278]. It would be unique if an isolated sugar was a substrate for a P450. However, despite exposing the protein to millimolar concentrations of L-arabinose, no spectral binding was observed and so confirmation of L-arabinose as a substrate was not possible. Finally, a range of amino acids was also tested. *C. jejuni* can use amino acids as a carbon source and amino acids are also used in the biosynthesis of vitamins [279]. Like all the other prospective substrates tested, no spectral binding was observed.

Because of the lack of success with the screening of typical P450 substrates that were easily accessible within the laboratory, a larger screen was undertaken using the FMP screening department in Berlin. Though a top ‘hit’ list of around 200 compounds was produced, some of these molecules were not suitable, as they either did not display typical absorption differences maxima and minima associated with substrate binding to P450s when tested in our laboratory, or else produced ‘noisy’ spectral data, suggestive of minor and/or irreproducible spectral perturbations. The decision to rank compounds based on the magnitude of the difference spectra that their binding (at near-saturating levels) produced did provide some successes, as hits 204073, 204074 and 213071 did produce good type I binding, which was also confirmed by EPR. The other proposed type I compounds selected failed to produce typical binding curves, despite promising difference spectra in the preliminary analyses. It was possible that absorption effects from certain compounds themselves affected the initial binding spectra, as even using dual beam UV-Visible spectroscopy failed to correct completely the absorption.
contributions from some of these compounds. EPR spectroscopy of the inhibitory ligands, (including econazole, 1-PIM and 204591) indicated that the ligated proteins existed in LS states, as expected. The g values observed were similar to those observed for the native protein, which might suggest that at least some of these inhibitors do not displace the 6th ligand water, but instead coordinate to the heme iron through the water ligand.

The redox potentials of the native (substrate-free) and substrate (204073) bound forms of the protein were analysed and heme iron midpoint potentials determined to be -418 mV and -379 mV, respectively, versus the SHE. The value for the native protein was very similar to other P450s, including the P450 BM3 heme domain (-427 mV) and M. tuberculosis CYP121 (-467 mV) [127, 267]. The full reduction of the substrate-free CYP172A1 may not have been achieved using dithionite, although spectral data from the Q-band region did indicate substantial conversion to the ferrous form. Thus, it is possible that the true heme iron potential is a little different from the value estimated. The binding of a substrate often increases the midpoint heme iron potential by ~+100 mV, favouring reduction of the heme iron. 204073 was apparently only able to cause a shift of + 40 mV. It is reasonable to suggest that this shift is not as large as would occur with a natural substrate, but 204073 still induced formation of a substantial amount of HS ferric heme iron. This modest effect of heme potential is consistent with data from the turnover experiments. Although 204073 was able to displace the axial ligand water and move the enzyme from LS to HS, no notable product formation was observed. By contrast, 213071 indicated much lower HS heme development when examined by UV-Visible and EPR spectroscopy, with only ~50 % conversion to HS. However, in this case there was clear evidence that it could be converted into product. The results from LC-MS analysis suggested that the sulphur atom on the 213071 might undergo single or double oxygen insertion chemistry to make sulfoxide and then, sulfinic acid or sulfenic acid derivatives. Sulphur atoms readily undergo oxidation reactions as they are relatively electron deficient, and sulfoxidation has been observed in reactions of BM3 and P450cam with sulphur containing substrates [289, 290]. The substrate conversion rate was not high, with only ~5 % conversion to the product mixture. This indicates that
although being able to bind to the P450 and produce a HS shift, 213071 was still not a very good substrate.

The turnover reactions also highlighted the ability for non native redox partners from *E. coli* and spinach to communicate with CYP172A1. The rate of Fe(II)CO complex formation was used to investigate the efficiency of electron transfer from these enzyme systems. When the protein was in a HS state (i.e. when bound to 204073) the rate of Fe(II)CO complex formation was faster than in the LS (substrate-free) state. This likely occurs as a direct consequence of the more positive ferric heme iron potential in the substrate-bound form, which is a control mechanism to ensure that the catalytic cycle is not initiated (with the potential to create reactive oxygen species) when a substrate is not bound for oxidation. CYP172A1 interaction with the spinach ferredoxin appeared to be preferred over the *E. coli* flavodoxin as the Fe(II)CO complex formation rate constant was greater with the ferredoxin compared to the flavodoxin (0.028 and 0.012 min\(^{-1}\), respectively) under identical conditions. Located within the *C. jejuni* genome are two ferredoxin proteins (FdxA and FdxB) that might be able to interact with and reduce CYP172A1 to facilitate its physiological role.

In this chapter a range of type II and type I ligands were characterised for binding to CYP172A1 by UV-Visible and EPR spectroscopy. A range of novel ligands that produced type I and type II heme Soret shifts were identified by compound screening and the HS inducing (substrate-like) compounds were further tested for product formation. Only a single compound, 213071, was shown to act as a substrate, with MS analysis indicating that oxidation of the 213071 sulphur occurred. Although not a physiologically relevant substrate, the result illustrates that protein does have catalytic function in oxygen activation and insertion into a substrate. Finally, it was demonstrated that CYP172A1 is readily reduced by non-native redox partners, and that these can support catalysis (although electron transfer kinetics are slow, likely underlying the modest substrate oxidation observed with the 213071 molecule). Of the two redox partners tested, it appears that CYP172A1 has a preference for the spinach ferredoxin over the *E. coli* flavodoxin. This may result from the more favourable potential of the
ferredoxin for electron transfer to the P450 (i.e. it has a more negative potential that the flavodoxin) as opposed to a higher affinity for CYP121 \textit{per se}. 
Chapter 6. Conclusions

6.1. Introduction
The work presented in this thesis concerns the characterisation of the lone P450 enzyme, CYP172A1, discovered in the genome of the pathogenic bacterium *C. jejuni*. This gene is also present in some *Campylobacter, Helicobacter* and *Comamonas* species. Prior to the start of this work, our knowledge and understanding of this unique enzyme was highly limited. The genomic location of the P450, along with previously published data, suggested that CYP172A1 may be involved in the production and/or regulation of the bacterial capsule. The *Campylobacter* species that lack the *cj1441c* gene also do not contain the *kps* genes, which might be suggestive that the P450 function is towards the bacterial capsule. The CPS is a structure on the outermost surface of *C. jejuni* which is important to the cell for serum resistance, adherence, invasion of epithelia and virulence [19]. However, due to a lack of distinct operons within this bacterium, genomic location is not sufficient for the assignment of gene product function, and as such investigations into the possible role of CYP172A1 were required. Attempts to characterise the P450 focussed on the use of recombinant protein expression as well as studying the functionality of the enzyme *in vivo*. The limited knowledge of CYP172A1 and its potentially important role in a pathogenic bacterium provided an interesting starting point for the research project.

6.2. Initial investigations into the role of CYP172A1 in *C. jejuni*
In order to study the role of the P450 *in vivo* it was first necessary to disrupt the gene expression and then to reintroduce a functional copy of the *cj1411c* gene into the bacterial chromosome. This was successfully achieved in two strains, *C. jejuni* 11168H and *C. jejuni* 81-176, as demonstrated using a CYP172A1-specific antibody. Initially a large amount of cross reactivity was observed on the immunoblots. However, this cross reactivity was reduced by premixing the rabbit antiserum with a P450 insertional knockout mutant of *C. jejuni*. The use of the recombinant CYP172A1 as a control also aided in correctly identifying the CYP172A1 band in the wild type and CYP172A1
complemented strains, while the relevant band was absent in the P450 insertional knockout mutant. This was an important result as it clearly identified that the targeted insertion of a kanamycin resistance cassette into the *cj1411c* gene resulted in the loss of the gene product, and that any differences in observed phenotypes between the wild type and P450 insertional knockout mutant could indicate a functional role for CYP172A1 in *C. jejuni*. The ability of the P450 deletion mutant to grow on agar plates and in culture also revealed that the P450 enzyme was not essential to the host for growth under these conditions, although this may not be the case during infection.

A range of techniques were employed to investigate the role of CYP172A1 in *C. jejuni*. The P450 insertional knockout mutants in both strains investigated had growth defects compared to the wild type strains. The mutant strains did not reach the same optical density in liquid culture compared to the wild type strains and this was partially rescued in the complemented strains. Initially it was thought that the differences in optical density might be due to different cell viabilities between the wild type and mutant strains. However, the colony forming units per ml were shown to be very similar for these strains. This led to the hypothesis that, if the wild type and mutant cell viabilities were actually similar, then perhaps the cellular structures might be altered between the wild type and P450 insertional knockout mutant resulting in a differing capacity to scatter light in the spectrophotometer. Using a combination of EM, AAG and motility testing it was demonstrated that a role for CYP172A1 in formation of the bacterial flagellum was highly unlikely, as P450 insertional knockout mutant strains were fully motile and images captured by EM revealed the presence of non-truncated polar flagella. Results from AAG experiments showed that the P450 insertional knockout mutant had a reduced ability to aggregate compared to the wild type and CYP172A1 complemented strains. However, the AAG was slower for a mutant that lacked a fully functional flagella, which again indicated that the role of CYP172A1 is not likely to be associated with the production of the flagellum. In electron micrographs of both the wild type and P450 insertional knockout mutant strains, cells of both strains displayed typical S shaped morphology, a common feature in *C. jejuni*. This finding was contradictory to a previously published report, which suggested that the loss of P450 function resulted in cells becoming “shorter and fatter” [202].
A broad range of antimicrobial agents were tested against the *C. jejuni* strains, in order to investigate if inactivation of the *cj1411c* gene affected the sensitivity towards these compounds. In most cases, the wild type and P450 insertional knockout mutant strains displayed identical MIC values. However, differences in sensitivities towards detergent-like compounds, including SDS and polymyxin B, were observed. These compounds target the bacterial surface membrane and, in the case of polymyxin B, the ability to cause cell lysis is affected by the presence of the LOS, and possibly to some extent by the integrity of the CPS as well. Further testing of polymyxin B confirmed consistent differences in sensitivities between the wild type and insertional knockout mutants using two different techniques. This further strengthened the hypothesis that CYP172A1 is involved in regulating the outer surface of the bacterium. The sensitivity of *C. jejuni* towards azole antifungal drugs was also investigated. These P450 inhibitor compounds were able to arrest cell growth, but in a P450 independent manner. Econazole, ketoconazole and miconazole were the most effective, with econazole preventing growth of the bacteria at a concentration of 2 mg/ml. The azoles used were all highly hydrophobic and were all prepared in DMSO. It is thus not fully clear whether the active concentration of these drugs was as high as that delivered, since precipitation of some of the azoles may have occurred in the agar used for these assays. Thus, it is possible their potency against *C. jejuni* is actually somewhat greater than the agar diffusion assays suggest.

Finally, gel based and spectroscopic techniques were used to investigate the LOS and CPS structures. Using silver staining, the LOS for all of the *C. jejuni* strains was visualised. The results indicated that no changes to the LOS structure occurred with the loss of CYP172A1 function. Despite repeated efforts, staining of *C. jejuni* CPS using Alcian Blue stain was not achieved. CPS from *C. jejuni* does not readily bind this stain, possibly due to a lack of charge on the CPS. Thus, a range of *C. jejuni* strains, which have previously demonstrated Alcian Blue staining, were tested, but again CPS was not detected. Due to the inability to visualise the CPS using the Alcian Blue stain, two spectroscopic techniques, FT-IR and SERS, were used to examine properties of the bacteria. FT-IR can characterise multiple components of bacterial samples, whereas SERS focuses primarily on the surface components. FT-IR was shown to be able to
resolve ten different *Campylobacter* species and strains, and the *C. jejuni* strains were separated from the other *Campylobacter* species using this method. A recent report, using a larger number of samples, obtained the same level of species distinction [222]. With both FT-IR and SERS the separation of the PE1-PE6 strains was not as might have been expected. The results possibly suggest that, while alterations to the CPS may have occurred in some of these strains, complete loss of the CPS does not occur. The capsular deficient mutant strain, ΔkpsM, and the PE2 strain did not overlap on any ordination plots. However, the separation of strains with a single gene alteration has not been previously attempted and it is possible that, at present, the resolution of the samples is not yet achievable using these methods.

6.3. Expression, purification and biophysical characterisation of CYP172A1
CYP172A1 was successfully purified following growth optimization in *E. coli* HMS174 (DE3) cells. The enzyme was soluble, but exhibited an unusual property. The enzyme was found to self-aggregate, which was unexpected as the majority of bacterial P450s exist as soluble monomers. This oligomerisation was first identified using size exclusion chromatography, as the majority of the P450 eluted with the void volume and eluted as a broad peak, rather than a single sharp peak as would be expected from a single monodispersed species. The polydispersity of CYP172A1 in solution was also above one, suggesting that aggregation was not a single stable complex, but rather involved multiple species, perhaps in a dynamic equilibrium. Upon SDS-PAGE analysis, a dominant band was observed at the expected M<sub>w</sub> of CYP172A1 (54.6 kDa). However, bands of higher M<sub>w</sub> were also consistently observed, and based on their apparent M<sub>w</sub> values were assumed to be P450 dimer and trimer species. When these bands were analysed by mass spectroscopy they were all identified as CYP172A1, based upon their peptide signatures. The reasoning behind the ability for CYP172A1 to self-aggregate remains unclear, but it may involve (at least in part) the formation of disulfide bridges. The addition of 100 mM BME reduced CYP172A1 self-aggregation and generated monomeric P450 fractions (observed by SDS-PAGE, size exclusion chromatography and MALLS). However, following the separation of the BME-treated protein by size
exclusion there was still a broad elution peak, indicating that disulfide bridge formation is not the only explanation for the phenomenon of self-aggregation of CYP172A1. A hydrophobicity plot did not reveal the presence of any obvious region(s) of amino acids that might form a nucleation point, and so it is unclear whether e.g. ionic or hydrophobic interactions are involved in the self-aggregation process.

Various methods were used in attempts to resolve the CYP172A1 self-aggregation, including removal of the His6 purification tag, and C and N terminal truncations of the P450 enzyme. The removal of the His6 tag indicated that it was not the driving force behind the P450 aggregation. Truncations at the N terminus of 38 and 109 amino acids, and a five amino acid truncation at the C terminus all failed to produce usable proteins. All three truncations resulted in the loss of heme incorporation compared to the full length CYP172A1. The C terminal and N-38 truncations resulted in lower heme, suggestive of a structural alteration to the protein that prevents heme binding. The N-109 truncation showed only partial heme incorporation, and still exhibited undesirable protein aggregation. The XtalPred software predicted that there might a region of disorder between amino acids 220 and 238 in CYP172A1, and so this region was compared against the sequences of structurally solved P450s. No major differences were observed, suggesting that the region is not disordered in CYP172A1 or in any of the structurally resolved P450s.

Although protein engineering did not provide truncated forms of CYP172A1 amenable to crystallization, biophysical characterisation of the intact CYP172A1 protein remained possible, and was to further provide information on spectroscopic and other properties of the P450. Importantly, the resting (ferric) CYP172A1 displays a UV-visible spectrum that is characteristic of most P450s, with a main Soret peak at 418 nm and distinct α, β and δ bands also present in the ferric LS state. Upon heme iron reduction with sodium dithionite and the addition of CO, the ferrous-CO P450 complex was formed, which further confirmed that CYP172A1 is indeed a true P450 protein. Homology modeling and far UV CD analysis also predicted that the P450 was predominantly made up of alpha helical secondary structure, again typical of the P450 class. The homology model built was based upon a P450 BM3 mutant, as the most similar structurally solved P450
based upon amino acid sequence identity. This was slightly surprising as CYP172A1 does not bind fatty acid substrates, while P450 BM3 shows high affinity for a range of fatty acids. The level of amino acid sequence identity between P450 BM3 heme domain and CYP172A1 is only ~30 %, and the sequence is most strongly retained around the conserved heme binding sequence that contains the conserved cysteine ligated to the heme iron.

6.4. Ligand binding to CYP172A1
A range of type II (inhibitor) and a few type I (substrate-like) ligands for CYP172A1 were identified by screening typical P450-binding molecules and by measuring changes in heme absorption in the UV-visible range of the CYP172A1 spectrum. Type I and type II interactions are identifiable from each other based upon the nature of the shift to the heme Soret peak, where type I compounds shift the peak to shorter wavelengths (~390 nm, resulting from displacement of the 6th ligand water on the heme iron) and type II compounds shift the peak to longer wavelengths, the extent of the shift is dependent on the ligating atom and the nature of the inhibitor. The best type II inhibitor compounds found for CYP172A1 were econazole and miconazole, from the antifungal class of compounds, with $K_d$’s in the nanomolar range. Type II binding was also observed with PIMs, fatty acid-linked imidazoles, imidazole and cyanide (Table 5.2). Both cyanide and imidazole had dissociation constants in the micromolar range. These compounds are both polar and their relatively poor binding probably indicates that, like most P450s, the active site of CYP172A1 is quite hydrophobic and disfavours the binding of hydrophilic molecules.

In order to identify type I, substrate-like molecules, CYP172A1 was subjected to compound screening at the FMP in Berlin using heme absorbance shift to identify hits. From the thousands of organic compounds screened, eight potential substrates were identified and ordered from ChemDiv for further studies. From the resulting eight optical titrations, only two of the compounds (204073 and 213071) were found to interact strongly with the P450. The remaining compounds produced strong absorbances in the 360-390 nm region in the absence of P450, probably explaining their selection in
the screening process. These results highlight the caution that must be applied in assigning binding of compounds to P450s by looking solely for the typical absorption troughs and peaks associated with general substrate-like binding to P450s. 204073 and 213071 share similar chemical structures, with the exception of the position of the branched benzene group (Figure 5.15). However, the position of this group appears important for the amount of HS heme iron conversion induced. This may be the result of 204073 binding in the substrate binding pocket more favourably than 213071, and allowing for a more efficient displacement of the distal water from the heme iron to move CYP172A1 into the HS state.

Following the identification of these substrate-like compounds, investigations into the thermodynamic properties of CYP172A1 were facilitated, and the heme iron midpoint potentials were determined in the presence and absence of substrate. Redox potentiometry using sodium dithionite (as a reductant) and potassium ferricyanide (as an oxidant) revealed that the heme Fe$^{3+}$/Fe$^{2+}$ midpoint potential of CYP172A1 was -420 mV vs. SHE. With the addition of the substrate 204073 to saturating levels, the midpoint potential positively shifted to -389 mV. This shift was not as large as has been seen for certain other bacterial P450s, which are often >100 mV. However, 204073 is not CYP172A1’s natural substrate and the binding of a true CYP172A1 substrate may prove more effective at inducing HS shift and a much larger shift in the heme iron potential.

Finally, it was shown that CYP172A1 has a catalytic function in converting the substrate 213071 into an oxidized product. This was seen as an oxidation on a sulfur atom in this substrate. Sulfoxidation reactions are not uncommon in P450-mediated reactions and sulfur atoms are good targets for oxidation. The turnover rate of CYP172A1 with 213071 was not fast, and the reaction did not go to completion. However, 213071 is not the natural substrate for CYP172A1, and the non-native redox partners used (typically *E. coli* FldR/FldA) are unlikely to drive P450 function as effectively as the *C. jejuni* redox partners. Collectively, these data likely point towards the types of substrate structure that CYP172A1 favours, but do not identify a specific physiological substrate.
6.5. Future directions

Further work on the function of CYP172A1 in C. jejuni should focus on the continuation of the investigations into the physiological role of this lone P450 enzyme in the bacterium. Although results obtained might suggest a possible function in CPS biosynthesis there is a lack of direct evidence. Future work should focus on renewed efforts to stain CPS with Alcian Blue dye. One approach might include examining other strains of C. jejuni with difference CPS structures that may more readily stain with this dye. The cj1411c gene could then be insertionally knocked out (as done in the current study) to allow for further investigations into the effects of cj1411c deletion and alterations to CPS structure. Structural analysis by mass spectrometry of the bacterial surface components would provide more detailed information to establish whether CYP172A1 has a role in production of LOS or CPS. It has been shown that the loss of either the CPS or LOS can affect the bacterium’s ability to successfully penetrate epithelial cells. Therefore, it would be of considerable interest to perform experiments in which wild type and P450 insertional knockout mutants of C. jejuni are incubated with epithelial cell lines, in order to compare their relative invasive capabilities and thus the relevance on the P450 activity for the infective process.

With regards to future studies on the recombinant expressed CYP172A1, further research efforts toward producing a P450 crystal structure should be made. In particular, a substrate-bound crystal structure would provide detailed understanding of the P450 active site topology and the interactions between the substrate and active site residues that are important in determining substrate selectivity. Prior to this being achievable, the polydisperse nature of CYP172A1 will need to be addressed. The use of gene expression systems with differing purification tags (e.g. GST or thioredoxin) might aid in reducing protein self aggregation. In addition use of non-typical buffers, reducing agents and ligands might stabilize the protein sufficiently to enable it to be crystallized. Mutations of the two peripheral (non-heme ligating) cysteine residues (e.g. to alanine) may also help to prevent the CYP172A1 self aggregation by preventing intermolecular disulfide bonds. The expression and purification of potential C. jejuni redox partners (e.g. ferredoxin and flavodoxin, and relevant reductases) could be undertaken to explore the rates of NAD(P)H-dependent electron transfer to CYP172A1, with the aim of
improving the electron transfer rate constants and identifying optimal redox partner combinations to improve efficiency of substrate turnover reactions. Further spectroscopic analysis of the CYP172A1 heme (e.g. using resonance Raman and magnetic CD methods) will also be important in establishing the properties on the heme centre; and combination of such approaches will ultimately lead to a detailed understanding of the structure of this unique P450 enzyme in this important human pathogen.

6.6. Concluding remarks
This thesis has been focused on improving our understanding of this under-studied P450, CYP172A1, in the food pathogen *C. jejuni*. From the experiments and assays that have been conducted it would appear that the function of the enzyme is likely to be associated with the biosynthesis or regulation of the *C. jejuni* capsular coat. This has been previously proposed due to the genomic location of the of the CYP172A1 gene (*cj1411c*) adjacent to a set of genes involved in capsular transport machinery (*kps* genes). The production of the P450 at both 37 °C and 42 °C suggests that the enzyme is not directly involved in pathogenicity but has a more general housekeeping role. The substrate for the P450 remains unknown and so the precise alteration and/or addition to the CPS cannot currently be characterized. Interestingly, CYP172A1 did not produce any HS heme binding effects upon the addition of fatty acids and so it might be possible that the enzyme’s primary role does not involve the lipid components of the outer membrane, the attachment point of the CPS moiety. Substrate-like molecules and azole inhibitors that bound in the nano-molar range contained hydrophobic groups and so it is possible to suggest that the natural substrate will contain a hydrophobic group similar to a purine based structure, which would allow for catalysis (most likely oxygen insertion) to occur. Although the sequence similarity of the primary sequence of CYP172A1 to other P450s was low, it does appear that CYP172A1 displays typical P450 structural and biophysical properties. The polydispersity of CYP172A1 is highly uncommon in soluble bacterial P450s and the aggregation appears to be mediated by multiple types of binding interactions, which may include hydrophobic and disulphide bridge bonds. It is interesting to note that current genomic annotations have genes *cj1410c-cj1412c* in a
small operon. *Cj1410c* and *cj1412c* are currently thought to be membrane associated and it is possible to speculate that CYP172A1 interacts with either of these two proteins, possibly via a hydrophobic interaction, which would go some way to explain the reasons for the polydispersity issues. By conducting the future experiments (Section 6.5) it is hoped that a P450-mediated modification to the CPS can be identified.
References


25. HPA. Factsheet on Cholera. 2008 [cited 2013 07.03].


References


95. SCOP. *Superfamily: Cytochrome P450*. 2009 [cited.


References


267


References


References


264. CSIRO. *Viewing and Scoring - Using CTWeb*. 2011 [cited 2013 05.03].


### Table S1. Primers used for PCR amplification.

Primers were used to check for genes and also to confirm the presence and the correctness of insertion knockouts. All sequences are given in 5’ to 3’ orientation. The numbers refer to those in the text. All primers were used at 20 pM in a PCR reaction.

<table>
<thead>
<tr>
<th>Primer number</th>
<th>Primer name</th>
<th>Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>375</td>
<td>Cj1411F</td>
<td>GCG TAG ATC TTG GCT TGA CGG</td>
</tr>
<tr>
<td>376</td>
<td>Cj1411R</td>
<td>TAC CAT TAG CCG AAC GCA CTG</td>
</tr>
<tr>
<td>680</td>
<td>Cj1411FlnkF</td>
<td>CGC ACG AGT GGC GTA AAT GC</td>
</tr>
<tr>
<td>681</td>
<td>Cj1411FlnkR</td>
<td>ACG CTC TCC CAA GCA TAG CC</td>
</tr>
<tr>
<td>1411</td>
<td>Cj1411CompF</td>
<td>GAC CCA TGG TAG TAT TAG TGT TAA AGG TGA TTA TG</td>
</tr>
<tr>
<td>1412</td>
<td>Cj1411CompR</td>
<td>ATG GCT AGC TAA AGT GCG TTT TAC ACG ACG</td>
</tr>
<tr>
<td>513</td>
<td>KanFout</td>
<td>CAT CCT CCT CGT CTT GGT AGC</td>
</tr>
<tr>
<td>561</td>
<td>KanRout</td>
<td>TTG CCT TCT GCG TCC GGT CG</td>
</tr>
<tr>
<td>311</td>
<td>Cat end seq F</td>
<td>GTG ACG GCT TTC ATG TTT GCC G</td>
</tr>
<tr>
<td>66</td>
<td>Cam 148R</td>
<td>GAT CAA AGC ATA ATC AAC CTG C</td>
</tr>
<tr>
<td>1</td>
<td>E109HF</td>
<td>G CTG CGC CCG TCG TTT CAT ATG ACG CGT ATC AAC</td>
</tr>
<tr>
<td>2</td>
<td>E109HR</td>
<td>GTT GAT ACG CGT CAT ATG AAA CGG GCG GCG CAG C</td>
</tr>
<tr>
<td>3</td>
<td>K38HF</td>
<td>GT CTG TAT GAA CGC TCA TAC CAT ATG CAG ACT GGC TAT GTG AA</td>
</tr>
<tr>
<td>4</td>
<td>K38HR</td>
<td>TT CAC ATA GCC AGT CTG CAT ATG GTA TGA GCG TTC ATA CAG AC</td>
</tr>
<tr>
<td>5</td>
<td>S449*F</td>
<td>CA AAT GGA ATG CGC ATT AAA TTT TAG AAA CGC AAA CTG TAA GGA TCC GG</td>
</tr>
<tr>
<td>6</td>
<td>S449*R</td>
<td>CC GGA TCC TTA CAG TTT GCG TTT CTA GAA TTT AAT GCG CAT TCC ATT TG</td>
</tr>
</tbody>
</table>
Figure S1. Schematic diagram of pGEMTEasy/cj1411c::kan plasmid. A plasmid map of pGEMTEasy/cj1411c::kan vector. The cj1411c gene (red) contains the insertion for the kanamycin resistance cassette (blue). BclI restriction sites are depicted and were used to insert the kanamycin resistance cassette.
Figure S2. Averaged SERS spectra acquired from *C. jejuni* strains. The spectra were created by averaging ~25 raw spectra collected for each cell type. Spectral peaks were observed ~750 cm\(^{-1}\) and between 1300-1600 cm\(^{-1}\). Cell 1 (PE4), cell 2 (PE5), cell 3 (PE6), cell 4 (PE1), cell 5 (PE2), cell 6 (PE3) and cell 7 (kpsM mutant). PE4 and PE5 appeared to have low signal to noise ratio. The average spectra acquired for PE6 was unique compared to the remaining samples.
Figure S3. Sensitivity of *C. jejuni* 81-176 to Polymyxin B containing agar plates. A) *C. jejuni* PE4 (left of red dotted line) and PE5 (right of dotted line) were spotted onto a blood agar plate which contains polymyxin B (8 µg/ml). Cells were serially diluted from $10^4$-$10^2$. B) *C. jejuni* 81176 PE4 (left of red dotted line) and PE5 (right of red dotted line) grown on a blood agar plate with no polymyxin B. Again cells are serially diluted from $10^4$-$10^2$. Cells were incubated for 48 hours.
Figure S4. Elution profile of CYP172A1 using size exclusion column chromatography. Two wavelengths were monitored, 280 nm (blue) and 420 nm (red). The protein eluted in two broad peaks, including at the void volume ~120 ml. The sample was run in 10 mM Tris buffer, with 100 mM KCl at pH 8.0.
Appendix

Figure S5. Mascot scores for proteins identified in Figure 4.6. Gel bands (red arrow on gel = A, lower blue = B and higher blue = C) were excised, trypsin digested and analysed by MS to examine peptide fragments. The fragments were compared the Mascot peptide sequence database. Matched fragments were ranked on their matches to proteins. Fragments in red indicate positive matches to the peptide fragments of the chosen protein. A score >90 indicates a strong match of the protein to a protein in the database.
<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidazole</td>
<td>8.0</td>
<td>57.0</td>
</tr>
<tr>
<td>Na/K phosphate</td>
<td>7.0</td>
<td>56.4</td>
</tr>
<tr>
<td>Bicine</td>
<td>8.5</td>
<td>54.8</td>
</tr>
<tr>
<td>EPPS</td>
<td>8.0</td>
<td>54.6</td>
</tr>
<tr>
<td>Tris</td>
<td>8.5</td>
<td>54.4</td>
</tr>
<tr>
<td>HEPES</td>
<td>7.5</td>
<td>54.2</td>
</tr>
<tr>
<td>Na/K phosphate</td>
<td>6.5</td>
<td>53.8</td>
</tr>
<tr>
<td>CHES</td>
<td>9.5</td>
<td>53.8</td>
</tr>
<tr>
<td>CHES</td>
<td>9.0</td>
<td>53.8</td>
</tr>
<tr>
<td>CAPS</td>
<td>10.0</td>
<td>53.6</td>
</tr>
<tr>
<td>Tris</td>
<td>7.5</td>
<td>53.2</td>
</tr>
<tr>
<td>MOPS</td>
<td>7.0</td>
<td>53.0</td>
</tr>
<tr>
<td>Ammonium acetate</td>
<td>7.0</td>
<td>52.8</td>
</tr>
<tr>
<td>ADA</td>
<td>6.5</td>
<td>52.4</td>
</tr>
<tr>
<td>Bis-Tris propane</td>
<td>6.5</td>
<td>50.4</td>
</tr>
<tr>
<td>MES</td>
<td>6.2</td>
<td>49.6</td>
</tr>
<tr>
<td>Bis-Tris</td>
<td>6.0</td>
<td>48.6</td>
</tr>
<tr>
<td>Sodium citrate</td>
<td>5.5</td>
<td>46.4</td>
</tr>
<tr>
<td>Na/K phosphate</td>
<td>5.0</td>
<td>37.8</td>
</tr>
<tr>
<td>PIPPS</td>
<td>3.7</td>
<td>29.4</td>
</tr>
<tr>
<td>Sodium acetate</td>
<td>4.5</td>
<td>28.8</td>
</tr>
<tr>
<td>Glycine</td>
<td>3.0</td>
<td>N/A</td>
</tr>
<tr>
<td>Citric acid</td>
<td>3.2 &amp; 4.0</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table S2. Calculated melting temperatures for CYP172A1 under different conditions using a JBS buffer screen. Buffers shown are from the JBS Solubility Kit, part A. Final concentrations for all buffers were 50 mM, diluted from 100 mM stocks. The buffers are listed in descending order of their effects on the stability of CYP172A1. N/A, indicates that the buffer did not produce a Tₘ for CYP172A1.
<table>
<thead>
<tr>
<th>Additive</th>
<th>Stock Concentration</th>
<th>$T_m$ (°C)</th>
<th>$T_m$ (°C) + 1-PIM&lt;sup&gt;a&lt;/sup&gt;</th>
<th>$T_m$ (°C) + Econazole&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>80 mM</td>
<td>57.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>200 mM</td>
<td>57.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>400 mM</td>
<td>57.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Glycerol</td>
<td>20 %</td>
<td>57.2</td>
<td>60.0</td>
<td>60.8</td>
</tr>
<tr>
<td></td>
<td>40 %</td>
<td>57.8</td>
<td>60.0</td>
<td>61.2</td>
</tr>
<tr>
<td>CHAPS</td>
<td>1.6 mM</td>
<td>56.0</td>
<td>58.2</td>
<td>59.8</td>
</tr>
<tr>
<td>Octyl glucopyranoside</td>
<td>0.4 %</td>
<td>56.0</td>
<td>58.2</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>4 %</td>
<td>56.2</td>
<td>58.2</td>
<td>60</td>
</tr>
<tr>
<td>Dodecyl maltoside</td>
<td>0.4 %</td>
<td>N/A</td>
<td>N/A</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>4 %</td>
<td>N/A</td>
<td>-</td>
<td>60.4</td>
</tr>
<tr>
<td>BME</td>
<td>40 mM</td>
<td>57.2</td>
<td>59.8</td>
<td>-</td>
</tr>
<tr>
<td>DTT</td>
<td>4 mM</td>
<td>57.0</td>
<td>59.6</td>
<td>60.4</td>
</tr>
<tr>
<td>TCEP</td>
<td>5 mM</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table S3. CYP172A1 melting temperatures in Na/K phosphate buffer with various additives and P450 ligands. Additives shown are from the JBS Solubility Kit, part B. Stock concentrations for all additives are shown. N/A indicates that the buffer did not produce a $T_m$ value for CYP172A1.<sup>a</sup> 1-PIM was at a concentration of 40 µM. <sup>b</sup> Econazole was at a concentration of 0.4 µM.
Figure S6. Elution profile of CYP172A1 from size exclusion column in the presence of 100 mM BME. Two wavelengths were monitored, 280 nm (blue) and 420 nm (red). The protein eluted in three broad peaks, including at the void volume ~8 ml. The sample was run in a 100 mM KP$_i$ buffer, with 100 mM KCl at pH 8.0 and in the presence of 100 mM BME.
Figure S7. Key features used by XtalPred prediction program in determining the likelihood of crystallisation of the CYP172A1 E109H truncation mutant. The software calculated the length (464 a.a.), pI (8.80), instability index (38.03), coiled regions (34 %), and disordered regions (17 a.a. long) and compared these parameters to all known proteins that have and have not been crystallised. These factors were combined to give a score of 3/5 for the probability of crystal formation in the CYP172A1 E109H mutant.
Figure S8. Key features used by XtalPred prediction program in determining the likelihood of crystallisation of the CYP172A1 K38H mutant. The software calculated the length (415 a.a.), pI (8.25), instability index (37.48), coiled regions (32 %), and disordered regions (18 a.a. long) and compared these parameters to all known proteins that have and have not been crystallised. These factors were combined to give a score of 3/5 for the probability of crystal formation in the CYP172A1 K38H mutant.
Figure S9. Key features used by XtalPred prediction program in determining the likelihood of crystallisation of the CYP172A1 S449* mutant. The software calculated the length (448 a.a.), pI (8.23), instability index (40.84), coiled regions (31 %), and disordered regions (18 a.a. long) and compared these parameters to all known proteins that have and have not been crystallised. These factors were combined to give a score of 3/5 for the probability of crystal formation for the CYP172A1 S449* mutant.
Figure S10. Western blots of CYP172A1 truncation mutants. Samples were run on 12% SDS-PAGE gels and transferred onto PVDF membranes. A) Immunodetection of the E109H truncation mutant. B) Immunodetection of the S449* truncation mutant and C) immunodetection of the K38H truncation mutant. A Western C marker was used to estimate protein sizes of 50 kDa (K38H), ~41 kDa (E109H) and ~54 kDa (S449*). S; soluble fraction, INS; insoluble fraction, NB; non-bound protein (to Ni-NTA column), W; wash from Ni-NTA column, integers indicate the concentration of imidazole (mM) used for CYP172A1 mutant elution from the Ni-NTA resin.
Figure S11. Optical titration of CYP172A1 with imidazole linked fatty acids, (A) ImC12, (B) ImC11 and (C) ImC10. Selected UV-Vis absorption spectra of CYP172A1 (3.9 µM) titrated with imidazole linked fatty acids. The resting Soret peak was at 418 nm and decreased in intensity and shifted to 421 nm with addition of the compounds. Binding curves from the titration, calculated by subtracting absorption values at the trough (414 nm) from those at the peak (432 nm) in each difference spectrum in order to obtain absorbance change values. The difference values were then plotted against the respective imidazole linked fatty acids and the data were fitted using equation 2 to generate $K_d$'s of 216.0 ± 8.1 µM, 148.8 ± 8.3 µM and 51.0 ± 4.4 µM for ImC10, ImC11 and ImC12, respectively.
Figure S12. Chemical structures of compounds ordered from ChemDiv stores. The 8 substrate-like compounds and an inhibitor-like compound (204591) that were purchased for further characterisation with CYP172A1.
Figure S13. Mass spectrum of 204073 substrate turnover reaction, mediated by CYP172A1. Following a turnover over reaction the mixture was extracted into DCM. The DCM was evaporated off and the residue dissolved into 100 μl MeOH:H₂O (50:50) prior to mass spectrometry. A single m/z ion was detected at 344, which corresponded to the size of the 204073 substrate.
Figure S14. Chemical structure of 213071 and potential products produced by turnover reactions. **A)** 213071 substrate with $M_w$ 334, **B)** sulfoxide formation with the addition of a single oxygen atom ($M_w$ 350), **C)** sulfinic acid formation with the insertion of two oxygen atoms ($M_w$ 368) and **D)** sulfenic acid formation ($M_w$ 353).