Infrared and UV-visible time-resolved techniques for the study of tetrapyrrole-based proteins

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

2013

Henry John Russell
Contents

List of Tables .................................................................................................................. 5
List of Figures .................................................................................................................. 6
Abstract .............................................................................................................................. 6
Declaration ......................................................................................................................... 10
Copyright Statement ......................................................................................................... 11
Acknowledgements ........................................................................................................... 12
Abbreviations and Symbols ............................................................................................... 13
Preface and Thesis Structure ............................................................................................. 16

1. Introduction .................................................................................................................... 19
   1.1 Spectroscopy Overview ............................................................................................. 19
   1.2 UV-visible and Infrared spectroscopy ....................................................................... 25
       1.2.1 Static Measurement .......................................................................................... 25
       1.2.1.1 UV-vis spectrophotometer ......................................................................... 25
       1.2.1.2 Fourier Transform Infrared Spectrometer .................................................... 27
       1.2.2 Rapid Mixing Techniques ................................................................................ 29
       1.2.2.1 Stopped-flow UV-vis spectroscopy ............................................................... 30
       1.2.2.2 Stopped-flow FT-IR spectroscopy ................................................................. 31
       1.2.3 Photoexcitation techniques .............................................................................. 32
       1.2.3.1 Laser-flash Photolysis ................................................................................ 33
       1.2.3.2 Ultrafast UV-vis Transient Absorption ......................................................... 36
       1.2.3.3 Time-resolved Infrared Spectroscopy .......................................................... 38
       1.2.4 Combination of techniques for protein study .................................................... 40
   1.3 Analysis of tetrapyrrole-based proteins by UV-vis and IR spectroscopy ..................... 41
       1.3.1 Cobalamins ....................................................................................................... 44
       1.3.1.1 B12-catalysis ............................................................................................. 45
       1.3.1.2 B12 photolysis ........................................................................................... 49
       1.3.1.3 Ethanolamine Ammonia Lyase ................................................................... 51
       1.3.2 Haems ............................................................................................................... 54
       1.3.2.1 Cytochrome c’ ........................................................................................... 58
   1.4 Thesis Aims and Objectives ...................................................................................... 63

2. Stopped-flow FT-IR Application Note: Temporal Resolution and Calibration of a
   Stopped-flow FT-IR Spectrometer using the Ester Hydrolysis of Methyl Chloroacetate ...... 64
   2.1 Introduction ............................................................................................................. 64
   2.2 Methodology ........................................................................................................... 64
   2.3 Experimental .......................................................................................................... 64
   2.4 Results and Conclusion ........................................................................................ 67

3. Ultrafast Infrared Spectral Fingerprints of Vitamin B12 and Related Cobalamins .. 68
   3.1 Abstract ................................................................................................................. 68
   3.2 Introduction ........................................................................................................... 69
   3.3 Materials and Methods ......................................................................................... 71
   3.4 Results and Discussion ......................................................................................... 72
       3.4.1 Adenosylcobalamin ....................................................................................... 73
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2</td>
<td>Methylcobalamin</td>
<td>77</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Cyanocobalamin and Hydroxocobalamin</td>
<td>80</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Model Frequency Calculations</td>
<td>82</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary and Conclusions</td>
<td>85</td>
</tr>
<tr>
<td>3.6</td>
<td>Supplementary Information</td>
<td>86</td>
</tr>
<tr>
<td>4.</td>
<td>Protein Motions Are Coupled to the Reaction Chemistry in Coenzyme B&lt;sub&gt;12&lt;/sub&gt;-Dependent Ethanolamine Ammonia Lyase</td>
<td>94</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>94</td>
</tr>
<tr>
<td>4.2</td>
<td>Materials and Methods</td>
<td>96</td>
</tr>
<tr>
<td>4.3</td>
<td>Results and Discussion</td>
<td>98</td>
</tr>
<tr>
<td>4.4</td>
<td>Summary and Conclusions</td>
<td>104</td>
</tr>
<tr>
<td>4.5</td>
<td>Supplementary Information</td>
<td>105</td>
</tr>
<tr>
<td>5.</td>
<td>Modulation of ligand-haem reactivity by binding pocket residues demonstrated in cytochrome c' over the femtosecond-second temporal range</td>
<td>113</td>
</tr>
<tr>
<td>5.1</td>
<td>Abstract</td>
<td>113</td>
</tr>
<tr>
<td>5.2</td>
<td>Introduction</td>
<td>113</td>
</tr>
<tr>
<td>5.3</td>
<td>Materials and Methods</td>
<td>117</td>
</tr>
<tr>
<td>5.4</td>
<td>Results and Discussion</td>
<td>120</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Wild-type AxCytcp</td>
<td>120</td>
</tr>
<tr>
<td>5.4.2</td>
<td>L16A AxCytcp</td>
<td>124</td>
</tr>
<tr>
<td>5.4.3</td>
<td>R124A AxCytcp</td>
<td>126</td>
</tr>
<tr>
<td>5.5</td>
<td>Concluding Remarks</td>
<td>132</td>
</tr>
<tr>
<td>5.6</td>
<td>Supplementary Information</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>Agreement of ΔA values between R124A TA and flash-photolysis experiments</td>
<td>142</td>
</tr>
<tr>
<td>6.</td>
<td>Conclusions and Outlook</td>
<td>143</td>
</tr>
<tr>
<td>6.1</td>
<td>Stopped-flow FT-IR method development</td>
<td>143</td>
</tr>
<tr>
<td>6.2</td>
<td>Ethanolamine Ammonia Lyase</td>
<td>143</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Implications for other AdoCbl-dependent enzymes</td>
<td>145</td>
</tr>
<tr>
<td>6.3</td>
<td>Cytochrome c'</td>
<td>146</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Implications for soluble guanylate cyclase</td>
<td>147</td>
</tr>
<tr>
<td>6.4</td>
<td>Final comments and future applications</td>
<td>148</td>
</tr>
<tr>
<td>7.</td>
<td>References</td>
<td>150</td>
</tr>
<tr>
<td>8.</td>
<td>Appendix</td>
<td>165</td>
</tr>
<tr>
<td>8.1</td>
<td>Ethanolamine Ammonia Lyase Overexpression / Purification Protocol</td>
<td>165</td>
</tr>
<tr>
<td>8.1.1</td>
<td>Protocol</td>
<td>165</td>
</tr>
<tr>
<td>8.1.2</td>
<td>Determination of sample purity</td>
<td>168</td>
</tr>
<tr>
<td>8.2</td>
<td>Steady-state assay of EAL</td>
<td>170</td>
</tr>
<tr>
<td>8.2.1</td>
<td>Reagent Preparation</td>
<td>170</td>
</tr>
<tr>
<td>8.2.2</td>
<td>Protocol</td>
<td>171</td>
</tr>
<tr>
<td>8.2.3</td>
<td>Activity Assay Standard Curve</td>
<td>172</td>
</tr>
<tr>
<td>8.3</td>
<td>&lt;sup&gt;15&lt;/sup&gt;N-isotopically labelled EAL protocol</td>
<td>173</td>
</tr>
<tr>
<td>8.3.1</td>
<td>Trial of minimal medium protocol</td>
<td>173</td>
</tr>
<tr>
<td>8.3.1.1</td>
<td>EAL activity trial using minimal medium</td>
<td>173</td>
</tr>
</tbody>
</table>
8.3.1.2 Determination of EAL yield and purity.........................................................174
8.3.2 15N-labelled EAL overexpression and purification........................................175
  8.3.2.1 Confirmation of successful 15N-labelling of EAL by mass spectrometry.....176
8.4 Generation of Cytochrome c’ L16A and R124A variants ................................179
  8.4.1 Plasmid design...............................................................................................179
  8.4.2 L16A and R124A mutagenesis......................................................................180
  8.4.3 Sequence data demonstrating successful mutagenesis..............................181
8.5 Cytochrome c’ Overexpression / Purification Protocol..................................183
  8.5.1 Protocol ........................................................................................................183
  8.5.2 Analysis of fractions.......................................................................................186

Word Count:  58384
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table 2.1</strong></td>
<td>Relationship between SF/FT-IR spectral resolution and acquisition rate</td>
<td>66</td>
</tr>
<tr>
<td><strong>Table 3.1</strong></td>
<td>AdoCbl TRIR and TA SVD principal kinetics</td>
<td>75</td>
</tr>
<tr>
<td><strong>Table 3.2</strong></td>
<td>MeCbl TRIR and TA SVD principal kinetics</td>
<td>79</td>
</tr>
<tr>
<td><strong>Table 3.3</strong></td>
<td>CNCbl and OHCbl TRIR and TA SVD principal kinetics</td>
<td>82</td>
</tr>
<tr>
<td><strong>Table S3.1</strong></td>
<td>CNCbl 375 and 525 nm excitation TA spectroscopy rate constants at various single wavelengths</td>
<td>89</td>
</tr>
<tr>
<td><strong>Table S3.2</strong></td>
<td>OHCbl 375 and 525 nm excitation TA spectroscopy rate constants at single wavelengths</td>
<td>91</td>
</tr>
<tr>
<td><strong>Table S4.1</strong></td>
<td>A comparison of the kinetic parameter from SVD analysis of both TRIR and TA data following photolysis of AdoCbl and EAI-AdoCbl</td>
<td>105</td>
</tr>
<tr>
<td><strong>Table 5.1</strong></td>
<td>Global fitting kinetic lifetimes of WT AxCytcp from TA and TRIR data</td>
<td>123</td>
</tr>
<tr>
<td><strong>Table 5.2</strong></td>
<td>Kinetic lifetime values from global fitting of TA and TRIR data for R124A across fs-µs (TA) and fs-ns (TRIR)</td>
<td>129</td>
</tr>
<tr>
<td><strong>Table S5.1</strong></td>
<td>Assignment of TRIR difference spectra feature for WT AxCytcp</td>
<td>136</td>
</tr>
<tr>
<td><strong>Table S5.2</strong></td>
<td>The influence of NO / protein concentration on L16A AxCytcp NO rebinding kinetics</td>
<td>139</td>
</tr>
<tr>
<td><strong>Table 8.1</strong></td>
<td>SDS-PAGE sample preparation for each of the EAL purification fractions</td>
<td>169</td>
</tr>
<tr>
<td><strong>Table 8.2</strong></td>
<td>Amino acids containing N atoms in their variable side chain</td>
<td>178</td>
</tr>
<tr>
<td><strong>Table 8.3</strong></td>
<td>Comparison of database peptide data with α-subunit mass spectrometry of $^{14}$N-EAL and $^{15}$N-EAL</td>
<td>178</td>
</tr>
<tr>
<td><strong>Table 8.4</strong></td>
<td>Comparison of database peptide data with β-subunit mass spectrometry of $^{14}$N-EAL and $^{15}$N-EAL</td>
<td>178</td>
</tr>
<tr>
<td><strong>Table 8.5</strong></td>
<td>Reagent volumes for mutagenesis of AxCytcp to L16A and R124A</td>
<td>180</td>
</tr>
</tbody>
</table>
### List of Figures

| Figure 1.1 | The electromagnetic (EM) spectrum | 20 |
| Figure 1.2 | The internal degrees of freedom of a diatomic molecule | 22 |
| Figure 1.3 | A Jablonski diagram for a hypothetical molecule | 24 |
| Figure 1.4 | Schematic of a standard FT-IR spectrometer | 28 |
| Figure 1.5 | Schematic of a standard SF/UV-vis spectrometer | 31 |
| Figure 1.6 | A Jablonski diagram representing photoexcitation of a hypothetical system | 33 |
| Figure 1.7 | Laser-flash photolysis instrumental layout | 34 |
| Figure 1.8 | TA spectroscopy instrument layout | 37 |
| Figure 1.9 | TRIR spectroscopy instrument layout | 38 |
| Figure 1.10 | Representation of protein chemical events and spectroscopic techniques available relative to time | 40 |
| Figure 1.11 | A simplified schematic of the tetrapyrrole biosynthetic pathway | 43 |
| Figure 1.12 | General mechanism of AdoCbl-dependent isomerases | 45 |
| Figure 1.13 | Reaction mechanism for AdoCbl-dependent ribonucleotide reductase | 47 |
| Figure 1.14 | General mechanism of MeCbl-dependent methyltransferases | 48 |
| Figure 1.15 | Role of Glu$^{287}$ during EAL catalysis | 54 |
| Figure 1.16 | The protein structural features that contribute to haem reactivity | 57 |
| Figure 1.17 | Overview of the NO/cGMP signalling pathway | 59 |
| Figure 1.18 | Haem structure of resting state, CO- and NO-bound AxCytcp | 60 |
| Figure 2.1 | Average SF/FT-IR difference spectra of MCA vs. NaOH relative to an early time point | 65 |
| Figure 2.2 | MCA vs. NaOH SF/FT-IR single frequency decays | 65 |
| Figure 2.3 | MCA vs. NaOH 1599 cm$^{-1}$ SF/FT-IR signal measured using various spectral resolutions | 66 |
| Figure 2.4 | MCA vs. NaOH 1599 cm$^{-1}$ SF/FT-IR signal measured using various shot volumes | 67 |
| Figure 2.5 | The relationship between SF/FT-IR shot volume and signal amplitude | 67 |
| Figure 3.1 | The chemical structure of cobalamin | 70 |
| Figure 3.2 | Ground state FT-IR spectra of AdoCbl, MeCbl, OHCbl and CNCbl | 72 |
| Figure 3.3 | Proposed model for the photolysis of AdoCbl | 74 |
| Figure 3.4 | AdoCbl static FT-IR spectrum; TRIR difference spectra; and SVD principal spectra | 75 |
| Figure 3.5 | AdoCbl TRIR SVD principal kinetics | 76 |
| Figure 3.6 | Proposed model for the photolysis of MeCbl | 78 |
| Figure 3.7 | MeCbl static FT-IR spectrum and TRIR difference spectra | 78 |
| Figure 3.8 | MeCbl TRIR SVD principal kinetic trace | 79 |
| Figure 3.9 | CNCbl static FT-IR spectrum and TRIR difference spectra | 81 |
| Figure 3.10 | OHCbl static FT-IR spectrum and TRIR difference spectra | 81 |
Figure 3.11  Cbl model structure used for DFT calculations  83
Figure 3.12  BP86 DFT calculated spectra for the Cbl model structure  84
Figure S3.1  Ground state FT-IR spectra for adenosine and AdoCbl  86
Figure S3.2  AdoCbl TRIR difference spectra and single frequency decays  86
Figure S3.3  AdoCbl TRIR single wavenumber traces at 1472 and 1551 cm\(^{-1}\)  87
Figure S3.4  Selected time point MeCbl TRIR difference spectra  87
Figure S3.5  Selected time point CNCbl TRIR difference spectra  88
Figure S3.6  CNCbl TA difference spectra with 375 nm photoexcitation  88
Figure S3.7  CNCbl TA difference spectra with 525 nm photoexcitation  89
Figure S3.8  Selected time point OHCbl TRIR difference spectra  90
Figure S3.9  OHCbl TA difference spectra with 375 nm photoexcitation  90
Figure S3.10  OHCbl TA difference spectra with 525 nm photoexcitation  91
Figure S3.11  Broad wavenumber BP86 DFT calculated spectra for the Cbl model structure  92
Figure S3.12  BP86/6-31G(d) DFT calculated spectra for the Cbl model structure  92
Figure S3.13  B3LYP/6-31G(d) DFT calculated spectra for the Cbl model structure  93
Figure 4.1  Current reaction mechanism for EAL  95
Figure 4.2  Comparison of AdoCbl and EAL static FT-IR spectra and TRIR difference spectra  99
Figure 4.3  SF/FT-IR analysis of EAL vs. ethylene glycol  100
Figure 4.4  Summary of SF/FT-IR analysis of EAL vs. ethanolamine, and EAL vs. 2-aminopropanol  102
Figure 4.5  SF/FT-IR difference spectra for \(^{15}\)N-labelled and non-labelled EAL vs. ethanolamine  103
Figure S4.1  EAL-AdoCbl TRIR SVD principal kinetic trace  105
Figure S4.2  SF/UV-vis measurement of EAL vs. ethylene glycol at 525 nm  106
Figure S4.3  Comparison of EAL TRIR and SF/FT-IR difference spectra  106
Figure S4.4  SF/UV-vis measurement of EAL vs. ethanolamine at 525 nm  107
Figure S4.5  SF/UV-vis measurement of EAL vs. 2-aminopropanol at 525 nm  107
Figure S4.6  SF/FT-IR difference spectra of EAL vs. ethanolamine between 1 and 2 seconds  108
Figure S4.7  Single wavenumber traces from SF/FT-IR difference spectra of EAL vs. ethanolamine at frequencies of major signal change  108
Figure S4.8  IR signal changes observed when acetaldehyde is left at room temperature for 10 minutes  109
Figure S4.9  SF/FTIR difference spectra of EAL vs. 2-aminopropanol between 5 and 10 seconds  109
Figure S4.10  Single wavenumber traces from SF/FT-IR difference spectra of EAL vs. 2-aminopropanol at frequencies of major signal change  110
Figure S4.11  Subsequent time point SF-FTIR difference spectra for both unlabelled and \(^{15}\)N-labelled EAL vs. ethanolamine  110
Figure 5.1  The haem binding pocket crystal structures of reduced and NO-bound AxCytcp
Figure 5.2  The currently proposed distal to proximal NO binding mechanism of AxCytcp
Figure 5.3  R124A AxCytcp mixture of NO-bound structures
Figure 5.4  NO-bound WT AxCytcp TA and TRIR difference spectra
Figure 5.5  NO-bound L16A AxCytcp TA and TRIR difference spectra
Figure 5.6  NO-bound R124A AxCytcp TA and TRIR difference spectra
Figure 5.7  UV-vis absorption difference spectra for NO-bound AxCytcp R124A from 2 ns – 10 µs
Figure 5.8  UV-vis 432 signal decay following photolysis of 5c-NO R124A from 1 ps – 100 ms.
Figure 5.9  The proposed geminate recombination mechanism for R124A
Figure S5.1  UV-vis spectra for reduced and NO-bound WT, R124A and L16A AxCytcp
Figure S5.2  FT-IR spectra for reduced and NO-bound WT, L16A and R124A AxCytcp, with difference spectra inset
Figure S5.3  Reduced WT AxCytcp TRIR difference spectra between 1-100 ps after excitation
Figure S5.4  Comparison of 1 ps TRIR difference spectra of WT AxCytcp with and without NO bound
Figure S5.5  Comparison of TRIR reduced samples for WT, R124A and L16A 1 ps spectra
Figure S5.6  Global fit of TA signal decay for NO-bound WT AxCytcp
Figure S5.7  The TA residual plot of WT AxCytcp NO rebinding kinetics following photoexcitation
Figure S5.8  1656 cm⁻¹ kinetic decay from WT TRIR data fitted to the sum of two exponentials.
Figure S5.9  Laser-flash photolysis data at 396 nm for NO-bound WT AxCytcp
Figure S5.10 Comparison of TRIR WT and L16A 1 ps difference spectra
Figure S5.11 Comparison of TRIR WT and R124A 1 ps difference spectra
Figure S5.12 Photodiode array spectra of R124A vs. 0.2 mM NO at time points between 1 – 500 ms post-mixing
Figure S5.13 NO concentration-dependence curve for R124A vs. NO acquired by SF/UV-vis
Figure S5.14  \( k_1 \) R124A laser-flash photolysis NO concentration-dependence curve
Figure S5.15  \( k_2 \) R124A laser-flash photolysis NO concentration-dependence curve
Figure 8.1  Example gel from SDS-PAGE of EAL purification fractions
Figure 8.2  EAL steady-state assay standard curve
Figure 8.3  SDS-PAGE of EAL purification using minimal medium protocol
Figure 8.4  SDS-PAGE of \( ^{15} \)N-labelled EAL overexpression and purification
Figure 8.5  \( \alpha \) subunit mass spectrometry spectra for \( ^{14} \)N-EAL and \( ^{15} \)N-EAL samples
<table>
<thead>
<tr>
<th>Figure 8.6</th>
<th>β subunit mass spectrometry spectra for $^{14}$N-EAL and $^{15}$N-EAL samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 8.7</td>
<td>UV-vis spectrum of high purity, as-isolated wild-type AxCytcp</td>
</tr>
<tr>
<td>Figure 8.8</td>
<td>UV-vis spectrum of high purity, as-isolated L16A AxCytcp</td>
</tr>
<tr>
<td>Figure 8.9</td>
<td>UV-vis spectrum of high purity, as-isolated R124A AxCytcp</td>
</tr>
</tbody>
</table>
Abstract

Infrared and UV-visible time-resolved techniques for the study of tetrapyrrole-based proteins

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Life Sciences by Henry J. Russell, 2013.

Tetrapyrrole-based proteins exhibit arguably the most diverse range of biological functions on the planet, facilitating photosynthesis, respiration and neurotransmission, amongst others. By virtue of their common biosynthetic pathway, all tetrapyrroles consist of a four pyrrole ring structure with a centrally chelated metal ion. This characteristic gives rise to absorbance of ultraviolet, visible and infrared radiation and thus allows their direct probing by ultraviolet-visible (UV-vis) and infrared (IR) spectroscopy. However, the surrounding protein structure can often remain spectroscopically silent in the UV-vis. In this thesis, the potential for combining time-resolved UV-vis with complementary IR spectroscopic techniques are presented, by study of two tetrapyrrole-based proteins; coenzyme B_{12}-dependent ethanolamine ammonia lyase (EAL) and haem-dependent cytochrome c'. These findings highlight the potential for probing chemical changes by UV-vis spectroscopy, whilst also identifying any coupled protein motions using IR spectroscopy.

Initially, a stopped-flow Fourier transform infrared (SF/FT-IR) instrument was characterised using test chemistry, which is summarised in an application note for use by the developer, TgK Scientific. Gaining greater insight of the SF/FT-IR instrument was essential for understanding the virtues of the technique prior to protein study. The first protein system analysed was EAL; a 5'-adenosylcobalamin (AdoCbl)-dependent eliminase. EAL mediates catalysis via AdoCbl cobalt-carbon bond cleavage, which yields a radical pair species that facilitates the reaction chemistry. The mechanism by which the protein scaffold guides radical pair formation and reactivity upon substrate binding is currently an area of debate. Using time-resolved IR techniques (including SF/FT-IR), in combination with UV-vis methods, real-time monitoring of protein dynamics during reaction has been achieved. The complementary IR and UV-vis signals suggest that protein dynamics guide cobalt-carbon bond cleavage, stabilisation of the adenosyl radical, and termination of turnover. Based on the EAL crystal structure, the contribution of a number of active site residues is discussed, in particular a mobile Glu_{287} residue that is thought to assist in the substrate trigger mechanism and radical pair stabilisation.

The second protein analysed was cytochrome c'; a NO binding haemoprotein with ligand binding properties analogous to the important eukaryotic signalling molecule soluble guanylate cyclase. Previous crystallographic and spectroscopic studies had implicated the importance of the haem binding pocket residues Leu_{16} and Arg_{124} during the NO binding mechanism in regulating ligand discrimination and haem stabilisation, respectively. By study of the cytochrome c' variants L16A and R124A in comparison with the wild-type using a range of UV-vis and IR photoexcitation techniques, the understanding of these residues' contribution to haem-NO reactivity has been furthered. In particular, Arg_{124} demonstrates protection against NO solvent escape, and is implicitly involved in haem-NO reactivity. The heightened importance of this Arg_{124} residue could have mechanistic implications for soluble guanylate cyclase, for which no crystal structure is available, and the protein motions coupled to catalysis remain under discussion.

These studies of EAL and cytochrome c' have not only extended their respective mechanistic understanding, but also demonstrated the power of coupling UV-vis and IR spectroscopy across wide time courses. By monitoring the electronic state of the chromophore by UV-vis, in concert with protein dynamics by infrared, this offers the opportunity to determine the methods by which proteins achieve their function.
Declaration

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 Statement

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.

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.

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

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on presentation of Theses.
Acknowledgements

There have been a great number of people who have helped me throughout my PhD, and I couldn’t have wished to be part of a more inclusive and caring group of lab members as over the past four years. To my supervisor Nigel, I want to thank him for providing me with the freedom to direct my PhD as I saw fit, and the guidance and willingness to help when things didn’t run smoothly. To Alex Jones, my adopted postdoc, special thanks must be given for his commitment to making me a well-rounded individual in every aspect of my work, and for the fun times we shared at both MIB and RAL. Sam Hay has also made a sizable contribution to my PhD, with his open door policy, happiness to help with data interpretation, and interesting stories of sporting injuries. Derren Heyes has been fantastic throughout my time at Manchester, whether a piece of equipment was broken, advice was required about planning an experiment, or work needed proof-reading, he would always drop everything to help. It’s just a shame he supports Liverpool, he’s in for a life of heartache. Since joining the group last year, Samantha Hardman has helped me a great deal in learning more about photoexcitation techniques and data analysis, and I feel a much more competent scientist for her input. Roger Kutta’s contribution to my PhD has been short in terms of time, but great in terms of input, and I am grateful for his constant willingness to extend my learning and understanding. Finally on a professional note, members of the lab who especially helped me settle in and understand general laboratory practice are Michi Sakuma and Jim Belcher, two of the friendliest people one is likely to meet.

On a personal note, the time doing my PhD has been a fantastic time in my life, not only because of the work I’ve done, but because of the friends and family who have been part of it. I’d like to thank in particular Laura Kemp; she’s been a great friend throughout my PhD and was always there when work didn’t go as planned. Finally, to my girlfriend Rosie, my parents Archie and Liz, and my sister Sally, thank you for all the love and support you’ve given to reach this point in my life.
## Abbreviations and Symbols

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>absorbance</td>
</tr>
<tr>
<td>Å</td>
<td>Ångström, (10^{-10}) m</td>
</tr>
<tr>
<td>A2P</td>
<td>1-amino-2-propanol</td>
</tr>
<tr>
<td>AdoCbl</td>
<td>5'-deoxyadenosylcobalamin</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>ATR</td>
<td>attenuated total reflectance</td>
</tr>
<tr>
<td>AxCytcp</td>
<td><em>Alcaligenes xylosoxidans</em> cytochrome c'</td>
</tr>
<tr>
<td>BLUF</td>
<td>blue light sensing using FAD</td>
</tr>
<tr>
<td>cAMP</td>
<td>cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>Cbl</td>
<td>cobalamin</td>
</tr>
<tr>
<td>cGK</td>
<td>cGMP-dependent protein kinases</td>
</tr>
<tr>
<td>CHES</td>
<td>N-Cyclohexyl-2-aminoethanesulfonic acid</td>
</tr>
<tr>
<td>cm^-1</td>
<td>wavenumbers</td>
</tr>
<tr>
<td>CM-cellulose</td>
<td>carboxymethyl-cellulose</td>
</tr>
<tr>
<td>CNCbl</td>
<td>cyanocobalamin</td>
</tr>
<tr>
<td>Co-C</td>
<td>cobalt-carbon bond</td>
</tr>
<tr>
<td>Co-N</td>
<td>cobalt-nitrogen bond</td>
</tr>
<tr>
<td>Co-R</td>
<td>cobalt-variable axial group bond</td>
</tr>
<tr>
<td>Cytcp</td>
<td>cytochrome c'</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DFT</td>
<td>density functional theory</td>
</tr>
<tr>
<td>DMB</td>
<td>5,6-dimethylbenzimidazole</td>
</tr>
<tr>
<td>DMB-on</td>
<td>cobalamin-binding protein with a intact DMB axial ligand</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTGS</td>
<td>deuterated triglycine sulphate</td>
</tr>
<tr>
<td>ε</td>
<td>extinction coefficient</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EAL</td>
<td>ethanolamine ammonia lyase</td>
</tr>
<tr>
<td>EM</td>
<td>electromagnetic</td>
</tr>
<tr>
<td>EPR</td>
<td>electron paramagnetic resonance</td>
</tr>
<tr>
<td>eut</td>
<td>ethanolamine utilisation operon</td>
</tr>
<tr>
<td>eV</td>
<td>electron volts</td>
</tr>
<tr>
<td>Fe^{2+} / Fe^{3+}</td>
<td>ferrous / ferric iron</td>
</tr>
<tr>
<td>FT-IR</td>
<td>Fourier transform infrared</td>
</tr>
<tr>
<td>GTP</td>
<td>guanine triphosphate</td>
</tr>
<tr>
<td>HEPES</td>
<td>2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid</td>
</tr>
<tr>
<td>His-on</td>
<td>cobalamin-binding protein with a histidine axial ligand</td>
</tr>
<tr>
<td>IR</td>
<td>infrared</td>
</tr>
<tr>
<td>J</td>
<td>Joule</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>$k$</td>
<td>rate constant</td>
</tr>
<tr>
<td>$K$</td>
<td>equilibrium constant</td>
</tr>
<tr>
<td>$k_{cat}$</td>
<td>catalytic rate constant</td>
</tr>
<tr>
<td>$K_d$</td>
<td>dissociation constant</td>
</tr>
<tr>
<td>$K_M$</td>
<td>Michaelis-Menten constant</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>wavelength</td>
</tr>
<tr>
<td>LMCT</td>
<td>ligand-to-metal charge-transfer</td>
</tr>
<tr>
<td>LOV</td>
<td>light, oxygen or voltage</td>
</tr>
<tr>
<td>M</td>
<td>molar, moles per litre</td>
</tr>
<tr>
<td>MCM</td>
<td>methylmalonyl-CoA mutase</td>
</tr>
<tr>
<td>MCT</td>
<td>mercury cadmium telluride</td>
</tr>
<tr>
<td>MeCbl</td>
<td>methylcobalamin</td>
</tr>
<tr>
<td>MES</td>
<td>2-(N-Morpholino)ethanesulfonic acid</td>
</tr>
<tr>
<td>MLCT</td>
<td>metal-to-ligand charge-transfer</td>
</tr>
<tr>
<td>Nd:YAG</td>
<td>neodymium-doped yttrium aluminum garnet</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>NOS</td>
<td>nitric oxide synthase</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>OHCbl</td>
<td>hydroxocobalamin</td>
</tr>
<tr>
<td>OPA</td>
<td>optical parametric amplifier</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>pD</td>
<td>corrected pH meter reading for solution in D$_2$O</td>
</tr>
<tr>
<td>PDE</td>
<td>phosphodiesterases</td>
</tr>
<tr>
<td>pET SEAL</td>
<td>pET vector containing EAL gene sequences</td>
</tr>
<tr>
<td>pH</td>
<td>negative logarithm of the proton (H$^+$) concentration</td>
</tr>
<tr>
<td>PMT</td>
<td>photomultiplier tube</td>
</tr>
<tr>
<td>RAL</td>
<td>Rutherford Appleton laboratories</td>
</tr>
<tr>
<td>SF/FT-IR</td>
<td>stopped-flow FT-IR</td>
</tr>
<tr>
<td>SF/UV-vis</td>
<td>stopped-flow UV-vis</td>
</tr>
<tr>
<td>sGC</td>
<td>soluble guanylate cyclase</td>
</tr>
<tr>
<td>SVD</td>
<td>singular value decomposition</td>
</tr>
<tr>
<td>$\tau$</td>
<td>time constant / life time</td>
</tr>
<tr>
<td>TA</td>
<td>ultrafast transient absorption (fs – µs)</td>
</tr>
<tr>
<td>TD-DFT</td>
<td>time-dependent density functional theory</td>
</tr>
<tr>
<td>TRIR</td>
<td>time-resolved infrared</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>UV-vis</td>
<td>ultraviolet-visible</td>
</tr>
<tr>
<td>$V_{max}$</td>
<td>maximum rate achieved by a Michaelis-Menten system</td>
</tr>
<tr>
<td>WLC</td>
<td>white light continuum</td>
</tr>
<tr>
<td>WT</td>
<td>wild-type AxCytcp</td>
</tr>
</tbody>
</table>
Amino Acids

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Three-letter code</th>
<th>Single-letter code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Ala</td>
<td>A</td>
</tr>
<tr>
<td>Arginine</td>
<td>Arg</td>
<td>R</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Asn</td>
<td>N</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>Asp</td>
<td>D</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Cys</td>
<td>C</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>Glu</td>
<td>E</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Gln</td>
<td>Q</td>
</tr>
<tr>
<td>Glycine</td>
<td>Gly</td>
<td>G</td>
</tr>
<tr>
<td>Histidine</td>
<td>His</td>
<td>H</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Ile</td>
<td>I</td>
</tr>
<tr>
<td>Leucine</td>
<td>Leu</td>
<td>L</td>
</tr>
<tr>
<td>Lysine</td>
<td>Lys</td>
<td>K</td>
</tr>
<tr>
<td>Methionine</td>
<td>Met</td>
<td>M</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>Phe</td>
<td>F</td>
</tr>
<tr>
<td>Proline</td>
<td>Pro</td>
<td>P</td>
</tr>
<tr>
<td>Serine</td>
<td>Ser</td>
<td>S</td>
</tr>
<tr>
<td>Threonine</td>
<td>Thr</td>
<td>T</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>Trp</td>
<td>W</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tyr</td>
<td>Y</td>
</tr>
<tr>
<td>Valine</td>
<td>Val</td>
<td>V</td>
</tr>
</tbody>
</table>
Preface and Thesis Structure

This thesis is presented in the “alternative format” style, whereby each of the results chapters presented are directly derived from, and presented in the style of, research publications. These publications (with the exception of the application note presented in Chapter 2) have been reported in peer-reviewed journals, and present the main findings of the PhD project. As a consequence of this layout, each of the results chapters include a targeted introduction relative to the research presented, in addition to materials and methods, results, discussion and summary sections. At the conclusion of each results chapter, where appropriate, supplementary information associated with the publication are also included. To allow continuity throughout the thesis, references are applied to the entire document and detailed in Chapter 7, with an increasing numbering style "(italic number)". All figures, tables and equations are numbered according to the chapter in which they are presented (e.g. for the fourth figure in Chapter 3 – Figure 4.3), with supplementary figures containing an “S” prefix. There is a general introductory chapter prior to the results chapters that outlines the aims and objectives of the thesis, in addition to appropriate information to the overall study. This introduction chapter provides the reader with a broad overview of with the current mechanistic understanding of each of the tetrapyrrole cofactors studied, as well as the chosen protein systems for analysis. Following the results chapters is a conclusions and outlook chapter that summarises the major findings of the research presented. This chapter also touches upon potential future research for the chosen protein systems, and the wider implications of the thesis conclusions. Since this publication-style thesis occasionally prohibits the illustration of general laboratory techniques and results crucial to a PhD project, a number of experimental protocols and findings are presented as appendices, which are appropriately referenced as footnotes throughout the thesis.

There are many benefits of presenting a thesis in the alternative format style. First, it encourages the candidate to focus on writing research papers during their PhD as opposed to concentrating on completion of the traditional thesis layout. This experience is invaluable in the development of a researcher, particularly one looking pursue a career in academia. Second, it presents research findings in a more concise manner, and in a style with which most academics are accustomed. In addition, these research publications have undergone peer-review, further enhancing their quality before presentation in the final thesis. Finally, the individual contributions of researchers associated with the PhD project are detailed more explicitly than in the traditional style, with many included as co-authors in the original publication. For the sake of clarity, the specific contributions of associated researchers are detailed below.
Thesis contributors

Nigel S. Scrutton oversaw the project progression in a supervisory role and has therefore contributed directly to every chapter of the thesis. Ted King, owner of TgK Scientific, provided technical expertise and instruction on the content desired for the application note presented in Chapter 2. My contribution for this application note included sourcing of an appropriate experiment for study, performing the experiments, and writing of the paper.

For the article presented in Chapter 3, Alex R. Jones and I were joint first name authors, since we contributed equally to the body of work, both experimentally and in construction of the manuscript. Gregory M. Greetham and Michael Towrie provided technical expertise and interpretation of data for the TRIR instrument at Rutherford Appleton Laboratories, while Sam Hay carried out computer simulation modelling analysis for the data produced.

For the communication in Chapter 4, I contributed the majority of experimental work and writing of the publication. Alex R. Jones assisted with project direction and interpretation of data, and aided the clarity of the publication by proof reading and amending when required. Gregory M. Greetham, Michael Towrie and Sam Hay’s contributions were the same as for Chapter 3.

For the article presented in Chapter 5, I implemented the experimental research and was responsible for writing of the article. Derren J. Heyes provided assistance with project direction, and Sam Hay performed data interpretation. Samantha J. O. Hardman provided technical expertise and interpretation of data for transient absorption experiments at the Photon Science Institute. Michael A. Hough provided plasmids for the protein variants and expertise on the system in a collaborative role. Gregory M. Greetham and Michael Towrie provided technical expertise and interpretation of data for the TRIR instrument.
List of publications


*These authors contributed equally


List of Application Notes

1. Introduction

This thesis covers the potential for coupling time-resolved infrared (IR) and UV-visible (UV-vis) spectroscopy techniques to solve long-standing mechanistic questions in protein systems. Time-resolved and static UV-vis spectroscopy techniques are highly established in biochemistry and have been developed to monitor the most elementary chemical events (on the fs time scale) through to whole domain motions (on the s time scale) in proteins. This is particularly beneficial when studying a protein with an associated organometallic cofactor, which results in rich spectral features that are dependent on the oxidation state of the metal, and the surrounding conjugated double bonds. However, despite this potential for directly monitoring protein cofactors, any associated protein motions often remain spectroscopically silent. By coupling these UV-vis spectroscopy techniques to equivalent IR measurements this offers the opportunity to measure reaction chemistry in accordance with related protein motions.

The introduction begins by providing an overview of the principles governing interaction of electromagnetic (EM) radiation with matter, highlighting a number of techniques that have harnessed these phenomena. UV-vis and IR spectroscopy are then covered in more detail, including time-resolved variations relevant to this thesis. Two protein systems with outstanding questions have been analysed using these methodologies; ethanolamine ammonia lyase (EAL) is a B$_{12}$-dependent eliminase that undertakes radical-mediated catalysis, whilst cytochrome c’ (Cytcp) is a gas-binding protein with unknown physiological function. These are covered in more detail in the latter part of the introduction, which highlights the past research, and potential for UV-vis and IR to increase their mechanistic understanding.

1.1 Spectroscopy Overview

The EM spectrum is the generic term for all possible frequencies of EM radiation. According to the wave-particle duality principle, the energy of a photon determines the frequency (and wavelength) of a wave and therefore all are equivalent measurements in spectroscopy ($E$). The relationship between photon energy, frequency and wavelength is outlined in Figure 1.1, along with the various classifications of EM radiation ($\lambda$). Due to the varying energies at different frequencies, the interaction of EM radiation with matter is altered accordingly. This is highlighted in the bottom panel of Figure 1.1, which details the resonating molecular energy levels at the respective classes of EM radiation. Various forms of molecular structure have energy levels, between which they can be excited by absorption of EM radiation, allowing the probing of molecules by spectroscopy. Nuclear spin-state energy levels arise when molecules containing nuclei of spin quantum number not equal to zero (due
to an odd number of protons and neutrons) are held in a constant magnetic field (3). The difference between these nuclear spin-state energy levels is relatively small, and therefore EM radiation in the “radiowave” spectral region is required to excite nuclei from one energy level to another. This phenomenon has been harnessed in nuclear magnetic resonance (NMR) spectroscopy, for which a number of approaches exist, and allow the mapping of molecular structure. Absorption of EM radiation can cause transitions between the spin-states of electrons in a similar manner to the spin state of magnetic nuclei (4). The magnetic field in which the electron is surrounded dictates the direction of its magnetic moment, and hence its spin vector, which can either be high energy \(m_s = \frac{1}{2}\) or low energy \(m_s = - \frac{1}{2}\). The difference between these energy levels resonates with microwave radiation, and this phenomenon has been exploited using the electron paramagnetic resonance (EPR) technique. Here, a constant microwave frequency is subjected to matter, then the magnetic field strength altered in order to probe the structure, reactivity and function of molecules.

**Figure 1.1. The electromagnetic (EM) spectrum.** Illustrated are various classified spectral regions, and equivalent wavelength, frequency, and energy of radiation (N.B. the rainbow spectral region represents the visible region). The bottom panel details the molecular structures that undergo excitation at resonating photon energies.
As the photon energy increases through the IR, visible and ultraviolet (UV) regions, it interacts with molecules in different ways. Before considering how the various photon energies affect molecular structure, it is important to consider intramolecular geometry. Each nucleus has three cartesian coordinates (x, y and z axes) and therefore a molecule of N atoms is considered to have 3N degrees of freedom (3). Thus, for a diatomic molecule this has six degrees of freedom as outlined in Figure 1.2. This includes three translational (movement of the molecule along an axis), two rotational (movement about an axis), and one vibrational (stretching of the chemical bond) motions. Translational and rotational motions constitute forms of Brownian motion of molecules provided by heat in the surrounding environment, and require relatively low levels of energy to occur. Therefore, as detailed in Figure 1.1, far-IR radiation is sufficient to affect the rotational and translational nature of a molecule. The only motion affecting the energy of the chemical bond is vibrational motion. For diatomic molecules the stretching mode is the only vibrational motion possible; however, with larger molecules there are a greater number of vibrational modes. The number of vibrational degrees of freedom for a molecule of N atoms is (3N-6) if non-linear and (3N-5) if linear; therefore, for a triatomic molecule this will consist of three vibrational modes if non-linear, and four vibrational modes if linear. Heterodiatomic molecules have a permanent dipole moment as a result of the differing electronegativities of the two atoms (5). These dipole moments oscillate at a certain frequency depending on the atomic centres bound together, but this normally falls within the range 200-4000 cm\(^{-1}\) (wavenumbers), i.e. the mid-IR range. Therefore, when incident radiation resonates with the frequency of the dipole this causes excitation of the molecular vibration. This phenomenon forms the basis of IR spectroscopy, which involves monitoring the absorption of IR radiation of varying frequencies by molecules. These absorption bands can be assigned to various vibrational modes and therefore give an indication of the arrangement of molecules (6). For this reason it is considered a form of vibrational spectroscopy. It is worth noting that homodimeric diatomic molecules are IR-silent since upon stretching there is no change in the dipole moment and therefore no energy absorbed. This explains the lack of IR absorption for both diatomic oxygen and nitrogen.
Figure 1.2. The internal degrees of freedom of a diatomic molecule. On the left is an outline of the possible 3 dimensions, through which a molecule can move, and in the centre are the various atomic motions of a diatomic molecule. There are three potential translational motions, two potential rotational, and one stretching vibration.

Increasing the photon energy to near-IR and UV-vis can affect the electronic configuration of molecules upon irradiation. This concerns the energy levels in which electrons reside upon excitation. In a thermally equilibrated system almost 100 % of molecules reside in the ground electronic state, as a result of the large energy difference between electronic levels. According to the Pauli Exclusion Principle, 2 electrons sharing an orbital must have antiparallel spins. Thus, one electron must have a spin of \( m_s = \frac{1}{2} \), and the other a spin of \( m_s = -\frac{1}{2} \). For most organic molecules the resulting overall multiplicity is 1 (\( M=2S+1=1 \), where \( S=0 \) is the vector sum of the single spin contribution), because the molecule exists with all its electrons paired in a closed shell configuration. The difference between electronic energy levels resonates with EM radiation of wavelength between 190 – 800 nm for the majority of molecules. Therefore, UV-vis radiation can be utilised to probe the electronic configuration of molecules by monitoring the absorbance of different frequencies of UV-vis. This is best illustrated with a Jablonski diagram as shown in Figure 1.3.
In a Jablonski diagram, the electronic energy levels for a molecule are represented using thick horizontal lines with energy increasing from low to high. Therefore, $S_0$, the lowest lying level, represents the ground state molecule, $S_1$ represents the first excited state, and $S_2$ represents the second excited state. Also included are the vibrational energy levels associated with each electronic excited state, with $v = 0, 1, 2, 3$ for $S_0$, and $v' = 0, 1, 2, 3$ for $S_1$, depicted by thin horizontal lines. It is worth noting that although electrons are unlikely to occupy the first excited state at thermal equilibrium, it is likely that several vibrational energy levels are occupied in the ground state. Upon absorption of a photon resonating with the energy gap between the ground state and an excited state (~ $10^{-16}$ s, or 0.1 fs), the molecule is excited to a vibrational level in an excited electronic state. Examples in Figure 1.3 are represented by purple and green arrows (7). The excited vibrational states (e.g. $v' = 1$ or 2) are unstable and relax to the ground vibrational state of $S_1$ ($v' = 0$) typically within 100 ps by non-radiative processes such as intermolecular collisions, in accordance with Kasha’s rule (3, 8). This method of non-radiative thermal relaxation is termed “internal conversion” (IC), indicated by wavy black arrows in Figure 1.3. Once in the $v' = 0$ state, the relaxation of the electron can occur via different processes. One method is to relax to the electronic ground state by IC, which includes an isoenergetic switch from the $S_1$ to the high vibrational state of $S_0$. This further results in release of the excess energy through thermal relaxation, for instance to the surrounding solvent. Another relaxation process involves emission of a photon (i.e. is radiative) equating to the difference between electronic levels, in a process termed fluorescence ($hv_f$ - yellow arrow). Both of these processes typically occur between 1 ps and 3 ns in solution.

Both IC and fluorescence relaxation processes typically involve singlet state relaxation to the ground state, however, the excited state configuration can convert isoenergetically to form the three vibrationally excited triplet states ($M=2S+1=3$, where $S=1$ is the vector sum of the single spin contribution) in a process termed intersystem crossing (ISC). Due to the very small energy gap between the three triplet states (a few cm$^{-1}$), in general all three can be considered as a single state ($T_1$). The energy gap between the $S_1$ and $T_1$ states is dictated by $2 \times$ the exchange integral ($2J$). As with singlet excited states, the vibrationally excited $T_1$ state also relaxes to the lowest vibrational level by IC in accordance with Kasha’s rule. Starting from $T_1$, relaxation to the ground state is forbidden on theoretical grounds according to the Pauli Exclusion Principle. Therefore this $T_1$ state can be long lived ($10^{-4} – 10$ s), and can either emit a photon in a process termed phosphorescence ($hv_p$), which requires a nuclear spin transition to maintain the overall spin multiplicity, or undergo ISC to reform a vibrationally excited $S_0$ which then can relax to the ground state via IC.
Figure 1.3. A Jablonski diagram for a hypothetical molecule. This outlines the absorption of a photon (purple and green arrows – $h\nu_A$), thermal (non-radiative) relaxation processes (internal conversion (IC) and intersystem crossing (ISC) – black arrows), fluorescence (yellow arrow – $h\nu_F$) and phosphorescence (red arrow – $h\nu_P$).

The processes of UV-vis absorption, fluorescence and phosphorescence can be exploited in electronic spectroscopy and used for characterisation of samples and their chemistry. UV-vis spectrometers are often used for identification of molecules by measuring spectra of absorbance against wavelength, or monitoring of absorbance change over time. Fluorometers on the other hand can be used to monitor fluorescence and phosphorescence of molecules. EM radiation of energy greater than UV has also been harnessed for molecular analysis. An example includes X-ray crystallography, where X-rays are directed towards a molecular crystal lattice, where it scatters, thereby producing a diffraction pattern (9). This diffraction pattern can be interpreted to produce electron density maps of molecules and therefore determine their structure. Another noteworthy scattering technique is Raman scattering, which involves the inelastic scattering of EM radiation in the near-IR, visible and near-UV range. This can be used to determine vibrational modes within molecules and is therefore another form of vibrational spectroscopy that yields complementary information to IR spectroscopy (10).
1.2 **UV-visible and Infrared spectroscopy**

UV-vis and IR spectroscopy exploit the phenomena of EM radiation interacting with matter, as covered in the previous section. In both cases, absorption of EM radiation can be used for identification of molecules based on characteristic spectra. For UV-vis spectroscopy, this returns spectra with broad electronic signals, whereas IR spectroscopy, by virtue of the narrower energy separation of vibrational bands, provides greater spectra detail. Both have static and time-resolved adaptations available, which will be covered in detail, with a focus on their application to the study of proteins.

1.2.1 **Static Measurement**

For both UV-vis and IR spectroscopy, spectrometers can be employed for static analysis of samples. In each case, a spectrum of absorbance is plotted against wavelength (wavenumber in the case of IR), with each technique employing different approaches for spectral acquisition.

1.2.1.1 **UV-vis spectrophotometer**

A typical UV-vis spectrophotometer generates a spectrum of absorbance against wavelength. The light source is often a tungsten halogen lamp for general purpose visible measurements, whereas high quality UV measurements (< 300 nm) are often provided by a deuterium lamp (11). This broadband light source is then converted into single wavelength light by a monochromator before interacting with the sample cell (normally a 1 cm pathlength cuvette), and detection of the transmitted light intensity measured by a photodetector. Absorbance is generated as a ratio of transmitted light \( I_{\text{out}} \) to incident light \( I_{\text{in}} \) by the equation \( A = \log \left( \frac{I_{\text{in}}}{I_{\text{out}}} \right) \), and therefore the instrument must be baselined before acquisition to remove any background absorbance arising from, for instance, the cuvette or buffer. One of the most common applications of this technique is determination of molecular concentration in solution, since concentration is directly proportional to absorbance. This is known as the Beer-Lambert law (Equation 1.1), which states that the absorbance at a certain wavelength \( A_{\lambda} \) can be related to the concentration of sample \( c, M \) and the pathlength of the cuvette \( l, \text{cm} \), by a constant termed the extinction coefficient \( \varepsilon_{\lambda}, \text{M}^{-1}\text{cm}^{-1} \).

\[
A_{\lambda} = \varepsilon_{\lambda} \times c \times l \tag{Eqn. 1.1}
\]

Another common function of UV-vis spectroscopy is the identification of molecules by their characteristic spectrum. Both of these experimental features lend themselves to studying proteins, which absorb strongly in the UV region at 220 nm as a result of peptide bonds.
connecting amino acids, and between 230-300 nm due to the aromatic amino acids phenylalanine, tryptophan and tyrosine (12). However, for UV-vis analysis it is most beneficial when proteins have associated metal ions or prosthetic groups with extended π-electron systems, as these absorb strongly in the visible region. These include chlorophyll, carotenoid, flavin, haem, cobalamin and retinal, amongst others. Absorption bands arising from these protein cofactors are sensitive to the surrounding polypeptide environment, the oxidation state of the chromophore, and the ligation state of the central metal. The effect of oxidation and ligation state on the absorbance of protein chromophores has been reported for a wide range of proteins, including extensive research on the famous haemoproteins haemoglobin and myoglobin (e.g. (13)). For ligand binding studies, this change in absorbance can be utilised to calculate the $K_d$ of protein-ligand binding, by titration of protein with increasing ligand concentration (12).

Another function of spectrophotometers is the measurement of absorbance at a single wavelength over a defined time course. This is particularly beneficial for characterisation of enzymes, for which steady-state turnover can be analysed, as well as identification of inhibitor molecules (14). Steady-state kinetics can be measured when the substrate of an enzyme is greatly in excess of the enzyme concentration. Steady-state analysis is possible on these timescales since the concentration of enzyme-bound intermediates has no net change (ignoring the initial mixing period) and therefore the reaction rate remains constant (15). UV-vis spectrometers can be used to calculate steady-state parameters of enzymes simply by manual mixing of an enzyme and its substrate in a cuvette of known pathlength. The absorbance change is then monitored at a specific wavelength, which reports on the degradation of substrate or evolution of product. For enzymes with substrates or products that give rise to UV-vis absorbance, this signal change can be monitored directly, whereas in the event of a spectroscopically-silent enzyme reaction this can be probed using a coupled assay. A coupled assay involves a secondary assay influenced by the enzyme reaction, of which the product causes a change in absorbance at a known wavelength. Using Michaelis-Menten kinetics calculations, the derived information from these assays can be used to determine the $K_m$ and $k_{cat}$ of an enzyme system (16, 17). This technique can be used to identify inhibitor molecules to enzymes by repeating this methodology with the addition of a suspected inhibitor. The effect addition of this molecule has on the values of $K_m$ and $V_{max}$ determines whether the molecule is an inhibitor to enzyme activity, and the inhibitor classification (18).
### 1.2.1.2 Fourier Transform Infrared Spectrometer

As indicated by the name, Fourier transform IR (FT-IR) spectrometers apply a Fourier transform to yield a spectrum of absorbance intensity against wavenumber, and the acquisition process is altogether different to UV-vis spectroscopy. Instead of generating monochromatic radiation (known as "dispersive measurement"), FT-IR spectrometers employ a broadband IR source that allows the measurement of a spectral window with a single acquisition (19). The basic outline of a standard FT-IR spectrometer is shown in Figure 1.4. Illustrated is an IR source, which in the case of most vibrational spectroscopy measurements is mid-IR (400 – 4000 cm\(^{-1}\)). This broadband radiation then enters a Michelson interferometer – the section of the spectrometer that facilitates non-dispersive acquisition (20). The interferometer contains a beamsplitter (usually KBr) angled at 45° to the incident IR radiation that reflects half of the radiation and transmitting the remaining half. After the beamsplitter, the IR beams reflect off a fixed or movable mirror, followed by recombination at the beamsplitter that results in interference of the beams. This process confers separation of the broadband IR radiation since constructive interference of the different frequency waves occurs at varying optical path differences (the measurement of distance the movable mirror has translocated from the central position). This then interacts with the sample held in either a transmission cell (with appropriate, non-absorbing windows such as calcium fluoride) of desired pathlength, or alternatively, for samples too thick or too strongly absorbing for transmission spectroscopy, attenuated total reflectance (ATR) cells can be used, which monitor the surface of materials (2). The IR absorbance is detected using either a mercury cadmium telluride (MCT) or deuterated triglycine sulphate (DTGS) detector. This results in an “interferogram”, which is a plot of detector response versus optical path difference (6). This can then be Fourier transformed to yield a spectrum of absorbance against wavenumber.

Both MCT and DTGS detector elements are transducers, converting infrared intensity into electrical signal, but operate in different ways (19). DTGS acts as a transducer by fluctuations in its temperature following exposure to IR radiation. This is a robust, inexpensive and simple method of detection; however, it requires a longer analysis time to generate spectra and therefore cannot be used for time-resolved measurements. Also, they are less sensitive than other detectors available. MCT is an alloy of three elements and is a semi-conductor, responding to IR radiation by excitation of electrons to a higher energy level. These excited electrons can respond to an applied voltage, which gives rise to an electrical current. Since the number of IR photons hitting the detector is directly proportional to the number of excited electrons, this can be used as a measure of IR intensity. This form of detection is very sensitive (10x more sensitive than DTGS), and can measure spectra very
quickly (therefore can be used for rapid scan measurement). The drawbacks of MCT detectors include their need to be cooled, which prevents heat from the detector giving rise to excited electrons, also MCT detectors can saturate easily, since too much IR hitting the detector will result in promotion of all available electrons to the higher energy level.

**Figure 1.4. Schematic of a standard FT-IR spectrometer.** This contains a Michelson interferometer in which a beamsplitter transfers half the mid-IR broadband radiation to a fixed mirror, and the other half to a movable mirror. The constructively interfered IR radiation then interacts with the sample and is detected, either by MCT or DTGS detector.

Clearly, FT-IR spectrometers are a useful tool for characterising molecules as a result of all heterodimeric molecules giving rise to IR absorption. FT-IR is a mostly non-destructive measurement device, ensuring the sample can be retained for further analysis. The versatility of the technique in terms of the samples possible for measurement also serves as an advantage, with gas, condensed, and solid state samples of small molecules, proteins, cells or tissues possible for biological study (21). Furthermore, very small quantities are required for analysis, with some sample cells as little as 5 µl is necessary for measurement (5). IR is particularly useful for study of protein functional groups that are spectroscopically silent in the UV-vis region. However, a drawback of FT-IR is the strong absorbance of certain solvents in spectroscopic window(s) of interest (21). One such solvent is H₂O, which absorbs strongly between 1350-1850 cm⁻¹ and 3400-3900 cm⁻¹; windows rich in spectral features including backbone amide regions and side chain vibrations of proteins (19, 22). This means that for H₂O to be used as a solvent, cell pathlengths should be restricted to below 10 µm (23). Alternatively, D₂O can be used as a H₂O substitute, as this renders a much reduced intensity of absorbance in these spectral regions (24, 25). However, use of D₂O as solvent prevents the
study of $^1$H/$^2$H exchange in the FT-IR spectra of proteins, a widely practiced and extremely valuable aid to band assignments (26).

Since the advent of FT-IR spectroscopy a large database for proteins and their individual amino acids has been subsequently developed (27). The technique can be used to measure protein structure, protein folding, unfolding and misfolding, and the molecular mechanism of protein reactions. When observing the IR spectrum of a protein the majority is contributed by the peptide backbone due to the abundance of repeating unit constituents. This gives rise to characteristic amide bands, which occur at general positions on the IR spectrum of proteins (28). Of these various amide bands, the amide I vibration is the most prominent, arising mainly as a result of the stretching of the C=O carbonyl bond in the peptide backbone, which gives a broad peak centred about 1650 cm$^{-1}$. Furthermore, the amide I band shape and position is altered as a result of the secondary structure of a folded protein and determination of protein secondary structure is currently the most common application of FT-IR spectroscopy for protein study (29, 30). Due to the sensitivity of the amide I band to protein secondary structure this allows analysis of protein folding, unfolding and aggregation to be analysed by FT-IR spectroscopy (31-33). Typically, while a fully folded protein exhibits a structured amide I band around 1650 cm$^{-1}$, an unfolded protein lacks absorbance in this region. Protein aggregates tend to show a peak value of 1620 cm$^{-1}$ or lower (6). In addition to the study of protein structure, the action of specific amino acid residues can be monitored by FT-IR spectroscopy since they give rise to characteristic peaks in an IR spectrum (27).

1.2.2 Rapid Mixing Techniques

Although static UV-vis and FT-IR offer a useful tool for identification and characterisation of molecules, and (in the case of UV-vis) a method by which to measure steady-state turnover of enzymes, there are a number of biomolecular events that remain unresolved. Time-resolved adaptations of UV-vis and IR spectroscopy provide the opportunity to fast dynamic processes. The essential feature of a time-resolved device is the use of a trigger method to initiate structural or chemical changes in a system, followed by the probing of these changes, often by spectroscopy (34). This can include perturbing the system with pressure, temperature, photoexcitation or mixing of reagents. Rapid mixing of reagents followed by simultaneous spectroscopic monitoring has proven to be an invaluable tool in this regard, and can be used with a range of spectroscopic probes (34). The most commonplace example of this practice is by use of stopped-flow apparatus, for which UV-vis and IR versions are available. Both use the same concept of mixing reagents, but have differing characteristics and will be covered in turn.
1.2.2.1  **Stopped-flow UV-vis spectroscopy**

Stopped-flow UV-vis (SF/UV-vis) spectroscopy is a highly accessible technique for the study of chemistry on a millisecond time scale. SF/UV-vis was first developed in 1923 for monitoring the velocity of fast reactions (35), and over the subsequent decades has become technologically refined (36), mainly due to advancements in the mixing apparatus. Stopped-flow involves placing reagents into individual syringes, which are then pneumatically driven into a mixing chamber followed by monitoring of spectroscopic changes in a transmission cell (37). The desired shot volume (normally ~ 300 µl) is set at the stop syringe, which upon filling simultaneously triggers acquisition (15). Detection can be set to monitor absorbance at a designated wavelength, which is generated using the monochromated output of a lamp (usually a xenon arc lamp), and the absorbance decay measured in real time by a photomultiplier tube (PMT). Alternatively, it is possible to collect time-courses at multiple wavelengths using non-monochromated lamp output (polychromatic radiation) and a photodiode array detector. This offers the opportunity to track spectral signatures of reactants and any intermediates formed more readily. However, this increased spectral detail is at the detriment of the overall data quality in terms of signal to noise, since the PMT is a more sensitive detection method than photodiode array. Furthermore, with photosensitive samples, photodiode array measurement can result in photobleaching due to increased light exposure. Modern SF/UV-vis instruments allow control over flow rate, shot volume, the ability to use a variety of transmission cells, and computational control over sample triggering and data acquisition. A schematic of the SF/UV-vis technique is outlined in Figure 1.5.

The time-resolution is limited by the mixing chamber and transport to the transmission cell, known as the “dead” time of the SF/UV-vis instrument, and is normally ~ 1 ms. The enhancement in temporal resolution over static measurement allows the real-time monitoring of a number of biomolecular processes including protein-protein interactions, ligand binding events, protein folding, enzyme reactions and chemical reactions, amongst others (38-40). Also, from the absorbance changes measured it is possible to determine reaction rates, the complexity of reaction mechanisms, and information on short-lived reaction intermediates. This can include pre-steady state measurements for enzyme kinetics, enabling the identification of reaction intermediates. SF/UV-vis is also useful for the analysis of protein-ligand interactions, since the formation of the protein-ligand complex often occurs over rapid time courses. The time resolution of this SF/UV-vis allows the $k_{on}$ and $k_{off}$ rates of ligand binding to be determined (41). For these purposes rapid-mixing techniques are invaluable, and can reveal information regarding protein dynamics such as changes in backbone conformation, hydrogen bonding, or orientation of side chains.
The limitation in the technique, as with UV-vis spectroscopy in general, is the reliance on a UV-vis absorbance signal for successful monitoring of reaction kinetics. For a number of enzyme and protein systems it is often possible to probe the role of protein cofactors during reaction, but more difficult to elucidate the role of the protein scaffold, which can be spectroscopically silent in the UV-vis. Therefore a complementary (or even alternative) method of measurement is the stopped-flow FT-IR (SF/FT-IR) technique.

Figure 1.5. Schematic of a standard SF/UV-vis spectrometer. Reagents are contained in syringes A and B and are then pneumatically driven into the transmission cell via a mixer (M). Here, UV-vis light is applied followed by detection of absorbance.

1.2.2.2 Stopped-flow FT-IR spectroscopy

SF/FT-IR operates on a similar principle to UV-vis, where reagents are loaded into two drive syringes, which are then driven into a sample mixing chamber before the sample cell. The sample cell is mounted directly in an FT-IR spectrometer, which allows signal acquisition. The stopped-flow drive is synchronised with the FT-IR spectrometer acquisition by way of an electrical signal. An initial limitation to the development of the technology was the high pressure required to ensure the two aqueous solutions flow rapidly through the 50 μm optical pathlength cuvette used for IR measurement (42). However, preliminary studies into such technology occurred 20 years ago (42), and over the intervening years an integrated mixer has been developed that is capable of mixing within 10 ms.

The ability to study proteins in their native state in aqueous solution or in membranes, with no requirement for immobilisation, labelling or other covalent modifications, makes SF/FT-IR a desirable technique. The time-resolution achieved with this
technique is modest in comparison with UV-vis stopped-flow (~ 70 ms per acquisition – more detail provided in Chapter 2), but still allows for the study of protein conformational changes, owing to the rapid-scan function of modern FT-IR spectrometers. The acquisition of whole spectra in rapid scan mode removes constraints with respect to the reversibility or reproducibility of the reaction studied, which is often a problem with step-scan or pulse-probe methods that acquire point by point (26). As a result of the vast number of vibrational modes in even small proteins \((10^4 \text{ for a 20 kDa protein (43)})\), spectroscopic changes arising from small structural changes can often be masked below large backbone absorptions. Therefore, for time-resolved FT-IR measurements of proteins, difference spectra are often generated relative to an early time point. This allows for easy identification of increasing and decreasing absorption bands during a reaction. In order to monitor these sometimes subtle changes, protein solutions must be of sufficiently high concentration to give large amide absorption bands, whilst avoiding total saturation of the IR signal. Unambiguous band assignments can be achieved by isotopic labelling of proteins causing shifts of bands to a different frequency, or by amino acid exchange, which can result in loss of spectral features (43). A number of SF/FT-IR publications of enzyme systems have been reported by Roger Thorneley et al. including establishing kinetic parameters in the pre-steady state and steady state (e.g. (44-46)), and band assignments using isotopically labelled protein or substrate (e.g. (44, 47)).

1.2.3 **Photoexcitation techniques**

An alternative trigger method for initiation of time-resolved measurements is photoexcitation of molecules. Unlike stopped-flow methods, this is able to achieve greater than 1 ms time resolution of acquisition. The method involves delivering a short pulse of light to perturb a system, followed by monitoring of spectroscopic signal changes. A Jablonski diagram representing photoexcitation of a hypothetical system is shown in Figure 1.6, which shows that photoexcitation of this molecule (or complex) results in an electronic transition between the ground state and the first or second excited state. These excited states can then be analysed using a second, probe beam. The time window following photoexcitation in which this monitoring occurs will dictate the identity of transient species probed, and therefore the ability to monitor a number of different spectroscopic time frames is a powerful tool. Laser-flash photolysis is a UV-vis spectroscopic technique with the ability to probe ns to 100’s of seconds after photoexcitation at specific wavelengths, therefore providing the opportunity to monitor large domain molecular movements, diffusion controlled reactions, triplet states, and singlet states. A complementary method is ultrafast UV-vis transient absorption (TA) spectroscopy, a pump-probe technique allowing probing on the fs – μs time
scale. This therefore covers more elementary chemical events such as proton and electron transfer, and allows the probing of photophysical events shortly after photoexcitation. Finally, time-resolved (TRIR) IR spectroscopy is a vibrational spectroscopy technique analogous to TA, but allowing the vibrational molecular state to be probed. The TRIR technique used in this thesis monitored from fs - ns.

![Jablonski Diagram](image)

**Figure 1.6. A Jablonski diagram representing photoexcitation of a hypothetical system.** Following excitation with a pump beam to the $S_1$ level, the electronic or vibrational state of the molecule is analysed using a probe beam.

Due to the strong absorption of visible light by associated metal ions and prosthetic groups of proteins, these chromophores provide a convenient means to both initiate and probe photochemistry. Since visible excitation also affects the vibrational level of proteins, these can be probed by TRIR in order to complement TA data. These photoexcitation techniques are also useful for the study of protein-ligand complexes, as the bond between the protein and ligand can be cleaved by laser excitation, thus creating a non-equilibrium state that can be monitored spectroscopically in real time.

### 1.2.3.1 Laser-flash Photolysis

This technique was developed over 50 years ago and represented a pioneering advancement in optical spectroscopy. Such was its importance, the researchers behind the development of laser-flash photolysis; Manfred Eigen, Ronald George Wreyford Norrish and George Porter, were jointly awarded the 1967 Nobel Prize in Chemistry (48, 49). The application of the technique was aided by advancements in laser technology during the 1950s.
which allowed the generation of short, coherent light pulses for excitation of samples, followed by measurement of relaxation dynamics. A schematic of the laser-flash photolysis instrument used to collect data for this thesis is outlined in Figure 1.7, and is representative of the general technique. In this case, the laser pulse is provided by a neodymium-doped yttrium aluminum garnet (Nd:YAG) laser operating at the second (532 nm) or third (355 nm) harmonic of the 1064 nm fundamental. This excites the sample contained within a cuvette, thereby perturbing the system. Any spectroscopic changes are monitored using an adapted spectrophotometer setup. Incident radiation is provided by a 150 W xenon arc lamp, which is pulsed for measurements < 1 ms, and adjusted to the probing wavelength by an input monochromator. After passing through the sample an output monochromator is required to remove any excitation light and retain the monochromated light from the xenon arc lamp, then the absorbance is measured using a fast PMT and converted to signal by a digital oscilloscope. A reasonable time resolution for the flash photolysis method is 1 – 10 ns, which is generally limited by the pulse width of the laser output (for Nd:YAG lasers this is typically ~ 8 ns). As a consequence of this setup, signal traces can only be acquired for a single wavelength per measurement. However, by measuring at a number of different wavelengths, differential absorbance (“action”) spectra can be generated, which provide a 2D data array of time-resolved UV-vis spectra.

Figure 1.7. Laser-flash photolysis instrumental layout. The sample is contained within the sample cuvette and excited by an Nd:YAG laser tuned to the desired excitation wavelength. The spectroscopic changes are then monitored at a chosen wavelength using a fast PMT.

Providing a short pulse of energy to a system, as in the case of laser-flash photolysis, has a number of applications. If the molecular structure is excited to a higher energy level followed by relaxation back to the ground state, any intermediates can be monitored by UV-vis, if they occur on the > ns timescale. This is particularly applicable if excitation yields intermediate singlet and/or triplet states (ns-s), or promotes rotational and translational motion (ns-µs). Aside from simple molecular excitation, light-induced conformational
changes in biological systems have broad application, particularly if those changes are part of its physiological role. Obvious examples include biological photoreceptors, which are suitable for study by flash photolysis as their mechanism can be triggered by a flash of laser light. These are signal transduction proteins, with the most important families including rhodopsins, phytochromes, xanthopsins, cryptochromes, phototrophins, and Blue Light sensing Using FAD (BLUF) proteins (52, 53). Each of these families functions via changes in chromophore configuration upon illumination, followed by triggering of a signal cascade. Some of the intermediates are extremely long lived, and therefore flash photolysis can be employed for their analysis (52). However, since a number of photochemical events occur on the < ns timescale, TA and TRIR techniques are often required for study of photoreceptors.

A particularly useful characteristic of protein-ligand complexes is the ability to photolytically cleave the weak sigma bond connecting the protein with ligand, resulting in ligand dissociation from the protein. Rebinding of the protein-ligand complex or competition between the ligand rebinding and that of another molecule can then be monitored. Another elegant method to quickly change the concentration of reactants is by releasing biologically active molecules from photolabile precursors (34, 54). This activity was first explored in 1977 by the release of cAMP into solution following photoexcitation (55), and these compounds were later described to be "caged" during a study that successfully yielded ATP (56). To date there are many established caged compounds including caged phosphate, GTP, ATP and Ca^{2+} amongst others. A popular caging element is p-hydroxyphenacyl (pHP-) as this releases its compound rapidly (sub-nanosecond) and does not proceed via intermediates (57). Another option is the use of ortho-alkylated nitrophenyl compounds, however, these are slower to release their caged compound as a result of intermediate formation (34). Caging allows the study of a number of enzyme and protein systems down to nanosecond time resolution, offering a great advantage over stopped-flow techniques. However, a drawback of caged compounds is the necessity for appropriate caged substrates for the enzyme system under study, which can sometimes be challenging or even impossible to achieve (26).

Although laser-flash photolysis greatly improves the temporal resolution with respect to rapid mixing methods, a number of chemical events remain unresolved. Since absorption of photons occur on the < fs timescale, events such as geminate recombination, vibrational motion, and weak bond cleavage are not detected, in addition to a number of photophysical electronic transitions and electron transfer processes. "Ultrafast" spectroscopic techniques offer the opportunity to monitor such chemical events.
1.2.3.2 Ultrafast UV-vis Transient Absorption

TA spectroscopy achieves fs time resolution by application of pump-probe technology. This allows the probing of molecular vibrations and elementary photochemical and photophysical events in real time including bond breaking, formation and structural rearrangements (58). Furthermore, the important processes of proton and electron transfer, as well as cis-trans isomerisations (rearrangements along a double bond) also occur on these timescales (59, 60). Pump-probe spectroscopy utilises a different excitation and detection method to laser-flash photolysis. Instead of the continuous monitoring of absorbance, the sample is first excited by a monochromatic pump beam of desired frequency, then the UV-vis spectrum is measured by a probe beam of broadband frequency after a certain time delay following excitation. This allows a “snapshot” spectrum to be acquired at defined time points after excitation, which if monitored at enough time delays, can return an overview of ultrafast events. A simplified schematic of the TA setup at the Photon Science Institute (University of Manchester) is outlined in Figure 1.8. Here, 3.5 W output radiation at 1 kHz repetition rate and ~ 100 fs pulse duration is produced by a Ti:sapphire amplifier system. The 800 nm amplifier output is then split, with half used for the probe beam, and half converted to the desired excitation wavelength by an optical parametric amplifier (OPA) to form the pump beam. The time delay after excitation is dictated by an optical delay line, via which the probe beam is diverted, followed by passage through a CaF₂ crystal, which generates the white light continuum (WLC). In order to remove the intense 800 nm radiation of the probe beam, the WLC passes through a filter before passage through a beamsplitter, with half used for measuring a reference beam and half for monitoring sample absorbance.

The TA instrument generates high signal to noise ratio absorbance difference spectra relative to the ground state species. Firstly, the reference detector monitors signal resulting from the WLC, which, when subtracted from the sample spectra, allows the reduction of laser shot-to-shot fluctuation. Secondly, the pump beam is passed through a chopper, which allows the removal of the excitation pump. This results in measurement of the sample ground state UV-vis spectrum followed by measurement of the excited state spectrum. Subtraction of the former from the latter generates a difference spectrum relative to the ground state species. It is worth noting that this pump-probe setup at Manchester (measuring 100 fs – 3.3 ns) is complemented by a setup capable of measuring signal changes from 1 ns – 400 μs, which also generates difference spectra. This uses the same pump beam and detectors as for the fs-ns setup; however, for this method the pump beam does not employ a chopper for baseline monitoring, and the probe beam is generated by a fibre based laser with spectral range 350 – 2400 nm.
As mentioned in the flash-photolysis section, photoreceptor proteins are readily probed by photoexcitation techniques, and this has often been exploited using the TA method \((61)\). In particular, photoreceptors containing BLUF or Light, Oxygen or Voltage (LOV) domains lend themselves to monitoring by UV-vis since they both contain a flavin cofactor, allowing a strict separation between the individual role of protein and cofactor \((52)\). This is in contrast to photoreceptor cofactors, which undergo isomerisation upon excitation, as these undergo overlapping isomerisation and twisting reactions in solution, making the mechanism of the protein difficult to assess. Many proteins (including photoreceptors) undertake electron, proton, or hydride transfer, which often occurs on the sub-picosecond timescale \((60)\). The advent of TA spectroscopy has allowed the probing of such chemical events, which previously were purely based on predictions. As with laser-flash photolysis, protein-ligand interactions can be probed with TA spectroscopy, but in this case the geminate recombination of ligand to protein can be analysed following photodissociation. However, unlike flash-photolysis, the use of caged compounds is often prohibited due to the time required for release of the caged molecule, and therefore probing of interactions between enzyme and substrate is often impossible on the ultrafast timescale \((52)\).

Figure 1.8. TA spectroscopy instrument layout. This schematic represents the setup at The University of Manchester, with WLC representing the white light continuum, and OPA the optical parametric amplifier.
1.2.3.3 Time-resolved Infrared Spectroscopy

TRIR spectroscopy represents a complementary technique to TA spectroscopy, and also employs a pump-probe setup with an optical excitation pulse, but instead utilises a broadband IR probe pulse for monitoring (62). This allows the analysis of vibrational transitions after photoexcitation on the fs to ns timescale. Since the earliest demonstration of TRIR over 25 years ago, which involved the study of excited state intramolecular hydrogen transfer (63), this technique has been developed and utilised in varying ways globally. Here, we focus specifically on the TRIR facility at Rutherford Appleton Laboratories (RAL) where a flexible, highly sensitive apparatus is available (64). A schematic of a typical TRIR experiment is shown in Figure 1.9, which produces an initial 10 kHz repetition rate, 1 mJ pulse energy, 800 nm, 50 fs output beam from a Ti:sapphire chirped pulse amplifier laser. This 800 nm incident radiation is then divided at a beamsplitter, with half used to contribute the UV-vis pump beam and the other half producing the IR probe beam. With regards to the pump beam, this travels via an optical delay line, which dictates the relative timing between pump and probe. Tunability of the pump and probe beams is achieved through femtosecond OPAs, which produce the pump wavelength desired (240 – 800 nm), whereas for the probe beam, difference frequency mixing is used to generate mid-IR output (2600 – 12000 nm).

Figure 1.9. TRIR spectroscopy instrument layout. The schematic above represents the instrument at Rutherford Appleton Laboratories (64) as described in text. Here, OPA represents the optical parametric amplifier, and DFG represents the difference frequency generator (DFG).
In order to minimise the laser shot-to-shot instability, the probe beam is passed via a beamsplitter with half directed towards a reference spectrometer and detector, and the other half transferred via the sample cell to an equivalent sample spectrometer and detector. This independent monitoring of probe spectral and intensity fluctuations simultaneously using an identical spectrometer allows for removal of much of the laser noise, generating higher quality data sets. As with TA measurements, a chopper is employed for the pump beam that allows acquisition of pump on, pump off spectra and therefore generation of difference spectra. The beams are focussed onto the sample to achieve typical beam diameters of 50 µm (probe) and 100 µm (pump). Samples are placed in Harrick cells, with UV – IR transmitting CaF₂ windows. The samples can be either as static or flowed systems, depending on the sample stability in the laser beam, and are rastered to avoid accumulation of photodegraded material on the cell windows. After the sample, the transmitted IR probe beam is collimated and focussed into two spectrographs; each with a 128-element photoconductive MCT detector, which, when used in parallel, can span a spectral window of > 500 cm⁻¹.

This TRIR setup generates spectra of resolution ~ 2 cm⁻¹, in addition to broad spectral windows (500 cm⁻¹). The latter is achieved due to shorter fs pulses generating broader probe spectra, which, when coupled with the 2 x 128-element probe arrays, allows numerous spectral features to be monitored simultaneously. This is particularly advantageous when using low volume or finite samples (such as when analysing biological species), as it allows for greater structural detail to be accrued, without having to record data in several spectral windows. TRIR acts as a complementary technique to TA spectroscopy for analysing proteins, since the same chemical events are monitored, but each provides different spectral information. One of the strengths of TRIR is the ability to gain deeper insights into the roles of nearby protein side chains in mediating function, particularly in the active site (62). This is especially beneficial for monitoring ligand dissociation events since the bond between protein and ligand can often be directly monitored, allowing greater insight into chemical events (e.g. (65)). Furthermore, returning to the inability of TA spectroscopy to monitor isomerisation events in some photosensory proteins, TRIR has demonstrated the ability to distinguish between protein and cofactor signals, providing deeper insight into photoreceptor function (e.g. (66)). As with other IR techniques, TRIR monitors vibrational changes following photoexcitation, which can include (amongst others) bond twisting, methyl group rotation, and local flexibility, all of which are on the ultrafast (fs-ns) timescale. A drawback of the technique is the need for high protein concentrations for analysis (~ mM) as a result of the narrow 10 – 100 µm transmission cell pathlength (62).
1.2.4 Combination of techniques for protein study

In protein systems containing organometallic chromophores such as tetrapyrroles, the combination of UV-vis and IR spectroscopic techniques offers great potential. In the case of UV-vis spectroscopy, this can be used for probing electronic transitions, and therefore provide insight into chemical transitions post-mixing or post-excitation, most likely arising from the protein cofactor. Whereas time-resolved IR spectroscopy can be used as a probe of protein dynamics that accompany the reaction chemistry, particularly by applying knowledge of when particular protein motions commonly occur.

As highlighted in the previous sections, functionally significant protein motions can span times ranging from fs – s (67). A summary of functionally important chemical and dynamic processes relative to their occurrence temporally is provided in Figure 1.10 (68). The lack of a defined end time for each of these events is intentional, since protein structure can often act to extend chemical lifetimes. Also contained within this image is the temporal range of the five time-resolved techniques available. In the case of UV-vis spectroscopy, these techniques offer time resolution over the entire range of time scales relevant for biomolecular dynamics. This offers a fantastic opportunity for study of proteins, provided they can be probed using photoexcitation techniques. Coupled with the vibrational spectroscopy techniques of TRIR (fs-ns) and SF/FT-IR (ms-s), additional information can be yielded regarding the inter- and intra-molecular bonding structure of proteins.

Figure 1.10. Representation of protein chemical events and spectroscopic techniques relative to time. Illustrated is a summary of a number of chemical events, dynamics and functions of proteins, compared with the temporal range of the five time-resolved techniques used in this thesis.
1.3 Analysis of tetrapyrrole-based proteins by UV-vis and IR spectroscopy

Tetrapyrroles are complex macrocycles that represent the most diverse array of cofactors to proteins, with roles in key biological processes such as electron transport, gas transport, light harvesting, respiration and catalysis (69-71). An important feature of tetrapyrroles is their ability to chelate diverse divalent metal ions, including Fe$^{2+}$, Mg$^{2+}$, Ni$^{2+}$ and Co$^{2+}$. The identity of the metal ion chelated depends entirely on the structure of the tetrapyrrole generated, and these are subdivided into haems, bacteriochlorophylls, chlorophylls, cobalamins, sirohaem, coenzyme F$\text{_{430}}$ and haem $d_1$ (71). Each of these groups contains four pyrrole rings linked through methylene bridges, with ring carbons having a variety of peripheral functional groups. The presence of metal ions, in addition to numerous conjugated double bonds within the macrocycle structure, results in strong absorption of UV-vis and mid-IR radiation. This is reflected in the array of colours between the various tetrapyrroles, leading to their consideration as “the pigments of life” (70). The biosynthesis of all tetrapyrroles begins with the formation of the small molecule 5-aminolevulinic acid, which acts as the source of all carbon and nitrogen atoms required for the formation of the tetrapyrrolic macrocycle. This molecule can be synthesised by one of two mechanisms – the C$_{5}$ pathway utilises the C$_{5}$-skeleton of glutamate (72) and can be found in most bacteria, all archaea and plants (73). Alternatively, the Shemin pathway, named after its discoverer, involves the condensation of succinyl-CoA and glycine followed by elimination of CO$_2$ (74). This mechanism is found in non-photosynthetic eukaryotes (including animals and fungi) and α-proteobacteria (75). Eight molecules of 5-aminolevulinic acid then undergo a three-step, enzyme catalysed process to form the first cyclic intermediate of the tetrapyrrole biosynthetic pathway, uroporphyrinogen III. Following the formation of uroporphyrinogen III, the biosynthetic pathways for generation of the various classes of tetrapyrroles diverge either immediately or subsequently, with varying degrees of modification.

A simplified version of the tetrapyrrole biosynthetic pathway is illustrated in Figure 1.11. This schematic is by no means exhaustive, with much greater detail offered elsewhere (e.g. (71)), but for our interests the key steps and products are presented. After formation of uroporphyrinogen III, four classes of tetrapyrroles are represented, which contain each of the four common divalent metal ions. The presented structure containing an iron (Fe$^{2+}$) is haem $b$, which has undergone the least modification from uroporphyrinogen III. Haem $b$ has been selected to represent the haem family since it is the most abundant haem in nature and acts as a precursor for a number of other haem types including haem $c$, $a$, and $o$ (76). Chlorophylls represent the tetrapyrrole group most closely associated to haems, and play a pivotal role in photosynthesis by trapping sunlight and converting this photon energy into chemical energy.
Presented in Figure 1.11 is the structure of chlorophyll \( a \), which is present in all oxygenic photosynthetic organisms as the primary photosynthetic pigment. Furthermore, chlorophyll \( a \) is hypothesised to act as the precursor to other chlorophyll isoforms via either direct (chlorophylls \( b \) and \( d \)) or indirect (chlorophyll \( c \)) modification processes (78). The structural similarity between haem and chlorophyll is reflected in the biosynthetic pathway, with the branching point between the two macrocycles coming three steps after uroporphyrinogen III, when protoporphyrin IX is chelated with \( \text{Mg}^{2+} \) as opposed to \( \text{Fe}^{2+} \) (haems). The noteworthy structural differences include the presence of an isocyclic ‘fifth’ ring, labelled E in Figure 1.11, in addition to a polyisoprene alcohol attached to ring D.

The \( \text{Ni}^{2+} \)-dependent tetrapyrrole coenzyme \( F_{430} \) is also presented, which is the only known nickel containing tetrapyrrole in nature (70). This porphyrin has been identified as the prosthetic group for methyl coenzyme M reductase of methanogenic archaea, an enzyme that catalyses the final step of methanogenesis. Additionally, coenzyme \( F_{430} \) is the most reduced tetrapyrrole in nature with only five conjugated double bonds, which results in the yellow colour of this porphyrin. The final structure illustrated in Figure 1.11 is that of cobalamin; the most sophisticated and elegant of all tetrapyrrole structures. This complexity is reflected in the necessity for 30 enzyme-catalysed steps during its biosynthesis, which restricts its metabolism to certain genera of bacteria and archaea (69, 79). Notable structural features include the ring contraction at the C20 position between pyrroles A and D, creating a “corrin” ring, in addition to the greater decoration around the periphery of the ring, with a number of propionamide and acetamide groups evident. Finally, unlike other tetrapyrroles, a lower axial ligand is provided to the central cobalt ion by a 5,6-dimethylbenzimidazole (DMB) moiety, and an upper axial ligand by a variable R group, which determines the identity of the cobalamin.

As a result of their respective interest to the biochemical community, haem and cobalamin-containing proteins have been selected for analysis by the experimental approaches outlined in Section 1.2. Both haems and cobalamins exhibit light sensitive coordination chemistry and return strong UV-vis and IR signals. Thus, their study harnesses the strengths of our time resolved spectroscopic techniques, and offers the opportunity to extend the mechanistic understanding of the chosen proteins. Haems and cobalamins will be covered in greater detail over the subsequent sections, in addition to an introduction to the protein systems selected for study.
Figure 1.11. A simplified schematic of the tetrapyrrole biosynthetic pathway. Illustrated are the terminal structures of four metabolic pathways, which represent the four known divalent metal ions. “A” represents an acetate group, “P” represents a propionate group.
1.3.1 Cobalamins

Cobalamin research began in the early 1920’s, when two physicians, Minot and Murphy, identified that pernicious anaemia could be cured with a diet containing whole liver (79). Pernicious anaemia was first identified in 1835, nearly 100 years before the cure was found, and therefore scientists were keen to identify the extrinsic factor contained within whole liver. This research continued for 20 years until two research groups at the powerful pharmaceutical firms Merck and Glaxo, isolated a red crystalline compound with the ability to treat pernicious anaemia, which they named vitamin B$_{12}$ (80). Structural determination of vitamin B$_{12}$ was initially attempted using traditional degradation techniques, which involved the breakdown of molecules followed by the analysis by spectroscopy and other methods of the various fractions. However, it soon became evident that this structure was one of extreme complexity in comparison with earlier molecules analysed by such methods, and it required the pioneering work of Dorothy Hodgkin and her laboratory to solve the structure of vitamin B$_{12}$. This was achieved in 1955 by X-ray crystallography, which revealed the structure illustrated in Figure 1.11, with a cyano (CN) group at the variable R position (81). However, it was soon revealed that this cobalamin (also called cyanocobalamin – CNCbl) was not biologically active, but rather a product of the extraction procedure from bacterial cultures (79).

The first biologically active cobalamin was crystallised in 1958, which acted as a coenzyme for glutamate mutase; an enzyme that catalyses the carbon skeleton rearrangement of glutamate to form a branched chain amino acid, β-methylaspartate (82). Due to its role in catalysis, this cobalamin was named coenzyme B$_{12}$. Once again, it was Hodgkin’s laboratory that solved the 3D structure of this compound by X-ray crystallography in 1961, revealing a structure of enhanced complexity to CNCbl (83). A 5’-deoxyadenosyl moiety was substituted at the upper axial position (thus, coenzyme B$_{12}$ is also named 5’-deoxyadenosylcobalamin – AdoCbl), which contained a cobalt-carbon (Co-C) bond previously unseen in nature, representing the first true organometallic system. The second, and only other, biological cobalamin was identified shortly afterwards. In 1962 a methylcobalamin (MeCbl) derivative was synthesised, which contained a methyl group as the upper axial ligand, and acted as a cofactor for methionine synthase in catalysing the conversion of homocysteine to methionine (84). This was confirmed by using an isotopically labelled [Me-$^{14}$C] MeCbl as the cofactor, which yielded methionine with $^{14}$C molecules incorporated into its methyl group.
1.3.1.1 \textit{B}_{12}\textsuperscript{-}catalysis

Since their initial discovery in the 1960’s, AdoCbl and MeCbl have been demonstrated to act as coenzyme and cofactor, respectively, to a number of important prokaryotic and eukaryotic enzymes (85). Both are utilised differently during catalysis and therefore will be considered in turn. All AdoCbl-dependent enzymes involve homolytic cleavage of the Co-C bond upon substrate binding to form a 5’-deoxyadenosyl / cob(II)alamin radical pair that mediates catalysis. The relative lability of the Co-C bond, with a predicted dissociation energy of \(~30\) kcal/mol, aids radical formation (86). Thus far only twelve AdoCbl-dependent enzymes have been identified, the majority of which carry out isomerisation reactions (87). Isomerisation involves the 1,2 interchange between a variable substituent and a hydrogen atom on adjacent carbons. These reactions often occur in bacterial strains, and play an important role in fermentation pathways. The general mechanism of AdoCbl-dependent isomerases is illustrated in Figure 1.12, where, following AdoCbl Co-C bond homolysis, the substrate is activated by abstraction of a relatively non-reactive H atom by the 5’-deoxyadenosyl radical. This step is common to all isomerases and yields a 5’-deoxyadenosine and substrate radical. The substrate radical undergoes 1,2 rearrangement of the variable group (X) to the adjacent carbon to form a product-like radical. The mechanistic details of this rearrangement step are dependent on the identity of X. Finally, the product-like radical abstracts an H atom from the methyl group of 5’-deoxyadenosine to form the isomerised product and the adenosyl radical, which immediately undergoes recombination with the Co\textsuperscript{2+} of cobalamin to reform AdoCbl.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure12.png}
\caption{General mechanism of AdoCbl-dependent isomerases. The migrating H atom is shown in red, whilst the migrating variable group (X) is shown in blue. X may be \(-\text{OH}, \text{NH}_3^+\) or a carbon-containing fragment.}
\end{figure}
There are currently three classes of AdoCbl-dependent isomerases, each of which are dependent on the identity of the migrating variable group. Class I comprise the AdoCbl-dependent carbon skeleton mutases that catalyse the interchange of hydrogen with a carbon-containing group, in doing so effecting the carbon skeletal arrangement. This is the largest class of isomerases with six members including glutamate mutase, 2-methyleneglutarate mutase, methylmalonyl-CoA mutase (MCM), isobutyryl-CoA mutase, hydroxybutyl-CoA mutase and ethylmalonyl-CoA mutase (88-94). It is worth noting that MCM is the only AdoCbl-dependent enzyme found in animals. X-ray structures of MCM and glutamate mutase, indicate that these enzymes bind AdoCbl in a “His-on” conformation, with the axially coordinated DMB buried into the core of the protein, and replaced by a conserved histidine residue (95-97). Furthermore, the substrate-bound structure of MCM demonstrated large scale motions in comparison to the non-substrate structure (96).

Class II AdoCbl-dependent isomerases are eliminases, which include diol dehydrase, glycerol dehydrase and ethanolamine ammonia lyase (EAL) (98). This subgroup catalyses the 1,2-rearrangements of either 1,2-diols to 1,1-diols or 1,2-amino-alcohols to 1,1-amino-alcohols, which then undergo enzyme-catalysed elimination of water or ammonia, respectively, to give aldehydes as the final product. Interestingly, X-ray structures of each of these enzymes reveal that they bind AdoCbl in a “DMB-on” conformation, and therefore retain the AdoCbl structure observed when the molecule is free in solution (99-101). The precise reason for why isomerases bind AdoCbl in various conformations is currently an area of debate. Finally, Class III AdoCbl-dependent isomerases are the aminomutases, for which only two enzymes are known, lysine 5,6-aminomutase and ornithine 4,5-aminomutase (102, 103). These catalyse the 1,2-migrations of amino groups, and require pyridoxal phosphate as a cofactor. As with the carbon skeleton mutases, X-ray structures have illustrated that these too bind AdoCbl in a “His-on” conformation, and also undergo large conformational changes upon substrate binding. Although the X-ray structures of a number of isomerases with and without substrate bound are available, the structure of the Michaelis complex, with the enzyme poised for catalysis, has remained elusive (87). Either the Co-C bond is cleaved, or the substrate is bound too far away from the cobalamin to give true insight into conformational changes coupled to catalysis.

The final AdoCbl-dependent enzyme, ribonucleotide reductase, is not contained in any of these categories, as it catalyses a reduction, not a rearrangement (104, 105). This class II ribonucleotide reductase is primarily found in microorganisms, and catalyses the conversion of ribonucleotide to deoxyribonucleotide; a fundamentally important process in DNA replication and repair (106, 107). The X-ray structure shows it binds AdoCbl in a “DMB-on” conformation similar to the eliminases (108). The currently accepted mechanism is
outlined in Figure 1.13 (adapted from [85]). As with all AdoCbl-dependent enzymes, the first step involves 5'-deoxyadenosyl / cob(II)alamin radical pair generation by Co-C bond cleavage. However, instead of interacting directly with the substrate, the 5'-deoxyadenosyl radical generates a reversible active site cysteiny radical that mediates the chemistry. This cysteiny radical abstracts the ribonucleotide 3'-H adjacent to the site of reduction, which activates the 2'-OH to become a good leaving group. Two active site cysteine residues donate two electrons to reduce the ribonucleotide, in doing so forming a disulphide bridge (which must be reduced with thioredoxin to enable multiple turnovers). Finally, the H atom is donated by the active site cysteine to form deoxyribonucleotide. Although this mechanism may appear distinct from the isomerases, the chemistry is similar to that for the 1,2 migration of −OH catalysed by diol dehydrase.

Figure 1.13. Reaction mechanism for AdoCbl-dependent ribonucleotide reductase. The migrating H atom is shown in red, with the eliminated −OH shown in blue.

MeCbl-dependent catalysis involves the transfer of a methyl group from a methyl donor to a methyl acceptor by so-called methyltransferases, which are involved in important metabolic pathways in humans, animals and bacteria ([85, 109-111]). These include methionine biosynthesis by methionine synthase, the metabolic formation of methane in methanogenic bacteria (thus contributing to their central pathway for energy production), and fixation of CO₂ via the acetyl-CoA pathway in anaerobic microbes ([109]). The role of MeCbl
in catalysis varies significantly from that of AdoCbl. As mentioned previously, MeCbl acts as a cofactor as opposed to a coenzyme during catalysis. All MeCbl-dependent methyltransferases follow the same general mechanism of catalysis, outlined in Figure 1.14, and bind MeCbl in a "His-on" conformation (112). Methyltransferases comprise three components, contained within different domains or polypeptide chains of the protein structure (111). The "B" component binds the methyl donor (X-CH\textsubscript{3}), the identity of which varies in molecular weight from small (e.g. methanol – 32 Da) to large (e.g. methyl tetrahydrofolate – 443 Da). As a result, the "B" component adopts appropriate structural features depending on the size of the methyl donor. Once bound, the methyl group is rendered electrophilic, normally by donation of a positive charge to the heteroatom attached to the methyl group. This allows the highly nucleophilic cob(I)alamin species held in the "C" component to abstract the methyl group from the methyl donor. The methyl acceptor (Y-) then binds to a cysteine-coordinated zinc ion within the "A" component. The role of the zinc is currently hypothesised to enhance the nucleophilicity of the thiolate methyl acceptor at neutral pH (113). This allows the heterolytic cleavage of MeCbl, and subsequent transfer of a methyl cation to the acceptor molecule, resulting in the formation of the final methylated product. Thus, in this mechanism MeCbl acts as an intermediate organometallic compound and undergoes heterolysis of the Co-C bond.

**Figure 1.14. General mechanism of MeCbl-dependent methyltransferases.** The methyl donor is depicted as -X, whilst the methyl acceptor is depicted as -Y. The various conformations of the protein are labelled in blue. The "B" component binds the methyl donor, the "C" component binds cobalamin, and the "A" component binds the methyl acceptor.
Catalysis mediated by AdoCbl and MeCbl plays a key role in a range of biological processes, with different modes of action. Furthermore, AdoCbl represents an excellent model system in demonstrating how proteins generate and control free radicals for catalytic function. One of the key mechanistic questions currently surrounding AdoCbl catalysis is determining the specific role of the protein scaffold in facilitating Co-C bond homolysis upon substrate binding. Indeed, this is relevant to EAL; the AdoCbl-dependent enzyme selected for study using our range of spectroscopic techniques, and is covered in greater depth in Section 1.3.1.3. However, prior to this, the complex photoresponse of cobalamins will be summarised, in addition to how this feature has been utilised to gain greater mechanistic insight into these complex molecules.

1.3.1.2 $B_{12}$ photolysis

The light sensitivity of AdoCbl was first identified in 1958, during the same study that outlined the crystallisation of the molecule by Barker et al. (82). It was observed that when a solution of coenzyme $B_{12}$ was illuminated with a tungsten lamp, there was a gradual colour change from yellow-orange to red, which was reflected in the altered UV-vis spectrum (82). It is now known that the chemical change observed upon exposure to light was the homolytic cleavage of the AdoCbl Co-C bond, yielding a 5'-deoxyadenosyl / cob(II)alamin radical pair similar to during catalysis. The eventual colour change observed is explained by the gradual conversion from AdoCbl to hydroxocobalamin (OHCbl), where a hydroxo group occupies the variable upper axial position, and occurs when AdoCbl is photolysed under aerobic conditions in an aqueous solution (114, 115). It is worth noting at this juncture, that after prolonged exposure (~ hours) to intense light, OHCbl also undergoes photolysis to yield a hydroxyl radical and cob(II)alamin (116). The fact that AdoCbl undergoes Co-C bond photolysis in a manner analogous to thermal Co-C bond homolysis during catalysis offers the opportunity to characterise radical pair dynamics without addition of substrate. Further advantages include the ability to suspend cobalamin molecules free in solution, allowing the study of (particularly in the case of AdoCbl and MeCbl) protein-bound and unbound cofactors. This is especially useful for the study of proteins that bind cobalamin in a "DMB-on" conformation since in this case the only chemical differences following photolysis can be assigned to protein-cofactor interactions.

AdoCbl has been studied extensively by TA spectroscopy, allowing the radical pair dynamics and photophysics following excitation to be characterised (117-121). These findings are covered in detail in Chapter 3 and therefore will not be considered here. It is worth mentioning that until recently the photolysis of AdoCbl was not thought to act mechanistically in nature, but rather a feature of the relative weakness of the Co-C bond.
However, recent studies have revealed that at least two organisms, *Myxococcus xanthus* and *Thermus thermophilus*, contain the putative photoreceptor CarH, which utilises AdoCbl as its chromophore (122-125). CarH is a transcription repressor, which, under dark conditions, oligomerises into a tetramer and binds to operator sequences, thereby blocking the downstream translation of carotenoids – proteins that protect cells against photooxidative damage (124). When exposed to light, AdoCbl bound to CarH undergoes homolytic cleavage, which leads to the dismantling of oligomeric structure and removal from the operator sequences.

MeCbl is also light-sensitive and has been the subject of a number of TA studies when free in solution (118, 120, 121, 126). In this case, when the Co-C bond (which incidentally is stronger than AdoCbl – 155 ± 12 kJ/mol vs. 134 kJ/mol (127-129)) photolyses, the cobalamin undergoes partial homolysis and formation of a metal-to-ligand charge transfer (MLCT) complex, the proportion of which is wavelength dependent. The MLCT complex can undergo subsequent homolysis, from which there is near complete methyl radical solvent escape, owing to the reactive nature of this species. With reference to the catalytic role of MeCbl in methyltransferases, where heterolytic cleavage is observed, study by photolysis does not offer a direct comparative study. However, there remains the opportunity to analyse the nature of the Co-C bond in MeCbl, in addition to the relationship between the methyl ligand and the main corrinoid structure. OHCbl and CNCbl, despite their lack of biological application, have also been studied by TA spectroscopy, mostly as a comparative study for AdoCbl and MeCbl (130, 131). In each case, upon pulsed photoexcitation, neither undergoes Co-R bond cleavage but instead illustrate corrin excitation signals that relax to the ground state within 100 ps, which reflects the increased bond strength in comparison with AdoCbl and MeCbl.

In the context of applying the techniques outlined in Section 1.2, the TA characteristics of AdoCbl, MeCbl, OHCbl and CNCbl have clearly been well documented, however, information regarding the vibrational signals after photolysis are fairly limited (with the exception of CNCbl, for which the Co-CN bond has been monitored (130)). Considering the number of vibrational modes contained within these cobalamin molecules, TRIR analysis of free cobalamins in solution offers the opportunity to monitor events following photolysis in greater detail with regards to the bonding and vibrational states of the molecules.
1.3.1.3 Ethanolamine Ammonia Lyase

EAL is an AdoCbl-dependent eliminase that is essential for growth on ethanolamine in the presence of coenzyme B$_{12}$ for a range of prokaryotes (98, 132-134). As with all AdoCbl-dependent isomerases, it employs radical-based catalysis and has outstanding questions regarding the reaction mechanism. EAL was first discovered by Bradbeer in 1965 as an essential component of choline fermenting Clostridium sp. (135, 136), and was later purified and characterised by Kaplan and Stadtman (137-139). These findings revealed that EAL catalysed the conversion of ethanolamine to the pre-product intermediate carbolamine via the typical 1,2-rearrangement reaction of AdoCbl-dependent isomerases. Carbolamine then undergoes enzyme-mediated dissociation into the final products acetaldehyde and ammonia (Equation 1.2).

\[
H_2NCH_2CH_2OH \xrightarrow{EAL} CH_3CHO + NH_3
\]  
(Eqn. 1.2)

Interest in this enzyme over the decades has been fuelled partly by a desire to understand how enzymes harness radical pair species to drive catalysis. In addition, EAL has medical implications since this and similar enzymes are functionally important in a number of food-borne pathogens. Therefore, understanding their structural and mechanistic function is key for engineering methods of inhibition in order to reduce food spoilage (101). EAL is encoded by genes contained within the eut (ethanolamine utilisation) operon, which contains genes encoding proteins responsible for ethanolamine degradation (140). eutB and eutC encode the large (α – 50 kDa) and small (β – 31 kDa) subunits, respectively, which form a dimer and bind one molecule of AdoCbl. This then forms a hexamer of dimers (α$_6$β$_6$), thereby resulting in six active sites per enzyme and a molecular mass of ~ 500 kDa (141, 142).

Due to the tendency for EAL to precipitate at the high concentrations required for X-ray crystallography, the 3D structure of the enzyme has required considerable time and effort (101). However, in 2010 it was reported that truncation of the β-subunit N-terminal resulted in significant enhancement of stability, with no apparent loss of enzyme activity (143). The structure of the truncated protein, in complex with a combination of coenzyme B$_{12}$ analogues and substrates, was reported later that year (101). As predicted by previous EPR (144) and amino acid sequencing (145) studies, EAL was shown to bind AdoCbl in a DMB-on conformation (101) and had an α$_6$β$_6$ quaternary structure. None of the structures reported contained the natural AdoCbl coenzyme, but instead a combination of CNCbl and adenylnpentylcobalamin (AdePeCbl – an inactive AdoCbl analog). These provided insights into the binding modes with and without ethanolamine or 2-aminopropanol bound.
Solving the 3D structure of EAL has allowed greater insights into predicting the mechanism by which the protein scaffold assists in cleaving the Co-C bond upon substrate binding. Previously, construction of the reaction mechanism focussed strongly on EPR findings, which monitored radical species (146-151), in conjunction with UV-vis spectroscopy, which followed the oxidation state of the AdoCbl cobalt during reaction (152, 153). SF/UV-vis in particular was useful for monitoring the pre-steady state kinetics for various substrates (153), in addition to kinetic isotope effect (KIE) studies, which revealed that Co-C bond cleavage is coupled to H-abstraction by the substrate (154). Although these EPR and UV-vis findings were useful in predicting a reaction mechanism (the mechanism based on these findings is illustrated in Figure 4.1), this focussed exclusively on the chemical state of AdoCbl, and failed to offer an explicit role for the protein during catalysis. A noteworthy exception prior to solution of the EAL crystal structure, were studies that revealed the critical role of Arg160 in EAL turnover (155). The crystal structure has assisted in interpreting the reaction mechanism in this regard, and provides evidence of a number of active site residues thought to interact with AdoCbl, in particular a mobile active site glutamate residue (Glu287) (101). Glu287 was clearly visible in the electron density map of substrate-bound structures, where it makes direct electrostatic contact with the ribose 2'-hydroxyl group of the 5'-deoxyadenosyl axial ligand, whereas it was almost invisible in the substrate-free structures (101).

In an effort to determine how the protein environment influences Co-C bond reactivity, a number of recent publications have employed photoexcitation techniques, which exploit AdoCbl homolysis upon light exposure. Since EAL binds AdoCbl in a DMB-on conformation, this provides the opportunity for photolysis measurements with and without protein bound, allowing direct comparisons to be made following Co-C bond cleavage. Similar studies of glutamate mutase by TA spectroscopy illustrated that the photochemistry of AdoCbl is indeed affected by its environment (156, 157). Magnetic field effect (MFE) continuous wave photolysis measurements of free and EAL-bound AdoCbl returned an enhanced MFE in the protein bound form, suggesting the AdoCbl-binding site acts as a cage, thus favouring geminate recombination (158). This was in contrast to MFE studies of EAL in which the Co-C bond was cleaved thermally by introduction of substrate. Here, no MFE was observed, this demonstrates limited geminate recombination owing to kinetic coupling between homolysis and subsequent H-abstraction from the substrate (159). Additional EAL photochemical findings include continuous wave measurements reported by Robertson et al. of holo-EAL and holo-EAL in complex with the inhibitor molecule 1-amino-2-propanol (A2P) over the 10^-7 – 10^-2 s timescale at room temperature (160). These measurements probed the separated radical pair dynamics of EAL, which decreased in quantum yield from 0.2 – 0.24 in
the case of free AdoCbl (117, 161), to 0.08 when bound to EAL, and 0.04 in the presence of A2P (160). Thus, it was concluded that the AdoCbl binding site suppresses photoprod
radical pair formation, whilst the substrate failed to induce radical pair stabilisation. Th
same group later reported complementary studies of the EAL ternary complex at cryogeni
temperatures (240 K), such that the complex was in a state of quasi-equilibrium, with a co
alamin / substrate radical pair (162). Under these conditions, no significant changes in s
or kinetics were observed upon substrate binding, leading to their assertion that there i
is a highly cooperative relationship between protein and Co-C bond cleavage. Such a c
operation would include formation of the equilibrium ternary complex, Co-C bond c
avage, and radical stabilisation, with protein dynamics guiding catalysis at every stage (162).

Finally, TA spectroscopic viscosity-dependence measurements of free AdoCbl in s
olution and in complex with EAL have been reported by our group (163). These findings w
ere in support of the continuous wave photolysis findings, and indicated that protein d
amics are coupled to Co-C bond homolysis, with an increase in viscosity leading to a d
crease in rate of radical pair separation. Given the timescales on which these rate v
ations occurred (ps-ns), and the reported crystal structure, the role of the Glu287 in th
-called “substrate trigger” mechanism was proposed as outlined in Figure 1.15 (163). Upon f
ormation of the ternary complex, Glu287 is mobilised towards the substrate, with el
rostatics leading to straining of the Co-C bond. The Co-C bond is then cleaved, with Glu287 c
ordinating ethanolamine via H-bonding of the amino and hydroxyl groups of the substra
t, and the newly formed 5’-deoxyadenosyl radical coming into direct contact with the a
stracted hydrogen atom. This fairly modest structural change upon substrate binding i
 contrast to that of other AdoCbl-dependent isomerases, and suggests a more subtle approa
towards Co-C bond homolysis.

Despite these recent developments in EAL mechanistic understanding, the protein c
tribution in real time has never been monitored. Considering the ability for SF-FTIR to d
ectly monitor post-mixing vibrational changes, this offers the opportunity to probe prot
structural changes following mixing with substrate. To aid interpretation of signals, these SF-
FTIR signals can be compared directly with equivalent stopped-flow UV-vis measurements,
which allow insight into the oxidation state of the cobalt at the beginning and end of turnover. I
addition, to complement published UV-vis photolysis data, TRIR can be used to probe any p
tein-AdoCbl interactions following Co-C bond photolysis. The combination of these t
ques can allow insights into the protein contribution during reaction and after radical p
pair formation.
Figure 1.15. Role of Glu\textsuperscript{287} during EAL catalysis. This scheme is based on (163), with the strained Co-C bond represented in red, and the migrating H atom represented in blue.

### 1.3.2 Haems

Haem is arguably the best known and most widespread prosthetic group in Nature, with diverse applications in various biological pathways (164). These responsibilities include, amongst others, simple electron transfer, oxygen transport and storage, catalysis, gas sensing, and transcription regulation (165). The variety of catalytic functions is extended further by combination of haem with additional cofactors, such as flavins, or the metal ions molybdenum or copper (166, 167). Haems were initially brought to the attention of the scientific community by the pioneering structural determination of myoglobin (168) and haemoglobin (169) in the late 1950’s. These oxygen storage and transport proteins account for the ubiquitous presence of haem throughout vertebrates and invertebrates, and have received continuous attention over the ensuing years. Myoglobin and haemoglobin contain a b-type haem (illustrated in Figure 1.11), which binds to proteins non-covalently. Of the different classes of haem molecules, haem b is certainly the most diverse in terms of application by proteins. In addition to oxygen storage and transport by the globins, b-type haems facilitate catalysis when bound to catalases, most peroxidases, b-type cytochromes and cytochromes P450 (CYP450s) (167). These enzymes exploit the redox properties of the haem prosthetic group to drive catalysis. In the case of the oxygenases (e.g. CYP450s and secondary amine monoxygenase), these extract an oxygen atom from O\textsubscript{2} to oxidise their substrate, whilst the peroxidases (e.g. horseradish peroxidase and catalases) carry out
substrate oxidation using H$_2$O$_2$ and other organoperoxidases. CYP450s in particular are crucial for mediating the oxidative metabolism of drugs and xenobiotics, and are found in all eukaryotic organisms (170, 171). The reactivity of the haem, and therefore its function, is governed by haem binding pocket architecture, with 20 different folding topologies currently recognised for binding of haem b (172).

The second major class of haem cofactor are c-type haems. In contrast to haem b, these are covalently linked to the haem binding pocket via thioether linkages that form between the thiol groups of cysteine residues on the protein and the 1'-carbons of the vinyl groups of porphyrin rings A and B (167). c-type haemoproteins are characterised by a CxxCH binding motif, referred to as the “fingerprint peptide”. The side of the haem containing this motif is termed the proximal face, with the opposing side named the distal face as a consequence of naming convention from the globin field (165, 173). The biological reasoning behind these thioether covalent linkages remains unclear, especially since haem c is metabolised from haem b, and therefore requires far greater biosynthetic energy input from the cell for production. Furthermore, haems b and c are similar in electronic structure and reduction potential, thus the rationale behind the presence of both isoforms is curious. c-type haems participate in a narrower range of biological processes that b-type haems, with many involved in simple electron transfer as cytochromes, which cycle between the reduced and oxidised oxidation states. They are, however, implicated in a wide range of energy transduction processes including photosynthesis, various respiratory pathways, and nitrogen cycling. The reduction potential of c-type haems also vary more greatly than that of haem b, spanning well over 1 V (165). The role of cytochromes in energy transduction processes is dictated by their reduction potential, as this determines their position along the potential gradient arising from all energy transduction processes.

In addition to their role in redox chemistry, c-type haemoproteins are also implicated in regulatory and signalling pathways. Haem sensor proteins bind an effector ligand (often nitric oxide (NO), carbon monoxide (CO) or molecular oxygen (O$_2$)) to the haem ferrous iron, followed by activation of a catalytic or DNA binding domain via protein conformational changes (174, 175). Examples include: CooA, a transcription factor that regulates CO metabolism in various bacteria (176); FixL, a transcription regulator of nitrogen fixation genes that senses low levels of O$_2$ intracellularly (177); and soluble guanylate cyclase (SGC), a NO sensor in higher organisms that regulates neurotransmission and vasodilation (178). In addition to the various roles of b and c type haemoproteins, vital roles in biochemistry are also implicated by other haem types. Noteworthy examples include haem a, which has a hydrophobic hydroxyethylfarnesyl chain on pyrrole ring A and a formyl group on ring D, and is utilised by cytochrome c oxidase, the terminal enzyme of most aerobic respiratory chains.
Also, haem $d_{1}$, which has two saturated pyrrole rings and two carbonyl groups, is bound to cytochrome $cd_{1}$, the enzyme that catalyses reduction of nitrite to nitric oxide in the periplasmic space of some bacteria (180).

Clearly, the identity of the haem subtype contributes to the function and reactivity of haemoproteins. However, the protein scaffold to which the haem is bound plays an equally important role in governing haem functionality. In the case of electron transfer proteins, structure modulates the reduction potential of the haem, whilst for ligand-binding proteins the scaffold can select between markedly similar diatomic molecules. Also, catalytic proteins utilise protein structure to guide the reactive haem towards substrate in a concerted manner.

Thus, ever since the X-ray crystallographic structure of myoglobin was solved (168), scientists have sought to fully understand how the protein scaffold dictates the properties of haem. It is currently thought that there are at least three major contributory factors at play. Firstly, the axial ligand(s) to the haem iron significantly affects the electronic state of the iron, and therefore its ability to withdraw electrons or bind ligands. Three axial ligands provided by the protein have been identified, including histidine (e.g. globins and peroxidases), tyrosine (catalases), and cysteine (chloroperoxidases and CYP450s) (173). It should be noted that hydrogen bonding to the iron proximal ligand also affects the distribution of charge and the strength of the ligand-metal bond, which contributes to haem reactivity (181-183).

Secondly, the hydrophobic environment and electrostatics affect the reduction potential of haem. If the polarity of the haem binding pocket environment increases this results in a decrease in the haem midpoint potential since these conditions stabilise a highly charged oxidised state (184). Therefore the proportion of hydrophobic to polar amino acids in the haem binding pocket directly influences the haem reduction potential. This feature affects haem-dependent redox proteins and enzymes most significantly, due to their reliance on a precise reduction potential for function. Finally, the solvent accessibility of the haem pocket has an influence on function, particularly for ligand binding by sensory and regulatory proteins. The degree of solvent exposure also affects the reduction potential of the haem, since upon greater solvent exposure this changes the polarity of the haem environment and therefore decreases the midpoint potential (185). An illustration highlighting the structural features that influence haem functionality is presented in Figure 1.16.
Figure 1.16. The protein structural features that contribute to haem reactivity. The reduction potential of haems are particularly affected by the axially bound ligand (His/Tyr/Cys in yellow), the proportion of polar and hydrophobic residues in the binding pocket shown in green and red, respectively, and the degree of solvent exposure, shown as H₂O molecules in blue.

In addition to influencing the reactivity of haem, protein structure also plays a critical role in the catalytic or DNA binding response of all haem-based signalling proteins. In this case, the conformational change arising from ligand binding results in stimulation of protein function, and understanding these structural changes represents an important field of biochemistry (173). Structural studies of these proteins have often focused on the haem binding domain alone, with the exception of CooA, for which the holo structure has been solved and offered great insights into coupling between domains upon ligand binding (186, 187). Given the electronic and vibrational signals arising from protein-haem bonding and coordination, time-resolved IR and UV-vis spectroscopic techniques lend themselves ideally for this study. For the photolysis methods (TA, flash-photolysis, TRIR), these are particularly useful for monitoring the influence of the haem pocket on ligand binding for sensory and regulatory proteins such as sGC, FixL and CooA, since the ligand-haem bond can be broken upon photoexcitation and the subsequent signals monitored. The question of how proteins are activated/deactivated upon binding of small molecules is one common to all regulatory
and sensory proteins. Therefore, the study of haem-based sensory proteins by spectroscopy has implications for understanding of non-haem proteins, with which they share a number of homologous features.

The haemoprotein selected for study by our range of spectroscopic techniques is cytochrome c' (Cytcp); a small NO binding protein with features analogous to sGC. The current understanding of this protein, and outstanding questions, will be covered in the subsequent section.

1.3.2.1 Cytochrome c'

Cytcp is a small, homodimeric protein found in the periplasmic space of a number of photosynthetic, denitrifying and nitrogen-fixing bacteria. Despite their extensive research over the past sixty years, the physiological role of Cytcp remains an area of debate, with current hypotheses suggesting a role in NO transport or reducing intracellular NO concentrations (188-191). Historically, Cytcp has remained the subject of scientific research due to its unusual ligand-binding and spectroscopic properties, in addition to its apparent similarity with sGC. As briefly covered in the previous section, sGC is a critical NO-signalling molecule that regulates important signalling pathways in higher organisms (192). This often occurs via the NO/cGMP-signalling pathway, which is initiated when a Ca\textsuperscript{2+}/calmodulin (CaM) complex binds to the enzyme nitric oxide synthase (NOS). NOS catalyses the oxidation of L-arginine (L-arg) and O\textsubscript{2} to L-citruline (L-cit) and NO. NO then diffuses across cell membranes to the target cell where it binds to a haem held in the β1 N-terminus subunit of sGC (193-195). sGC exists as a heterodimer, with the α1β1 isoform the most widespread in the cytosol (196). Upon binding of NO, the catalytic domain, which is currently thought to be controlled by residues in the C-termini of the α1 and β1 subunits (197), converts GTP to the second messenger cGMP and pyrophosphate (PPi). cGMP then binds to any of cGMP-dependent protein kinases (cGK), phosphodiesterases (PDE), or cGMP-gated ion channels, which regulate physiological functions including vasodilation, neurotransmission and platelet aggregation. The described NO/cGMP signalling pathway is outlined in Figure 1.17. The study of sGC has been hindered due to the current limitations in protein yield via bacterial expression systems. These techniques often result in insoluble sGC fractions or low protein yield, which has stifled X-ray crystallographic analyses (192). Therefore a great deal of interpretation of sGC function, particularly the molecular mechanism of how sGC is turned on/off upon binding of NO, has been provided by study of homologous systems such as Cytcp. The thorough understanding of sGC is especially important from a therapeutic viewpoint, since a number of diseases arise from defects in the NO/cGMP signalling pathway.
Figure 1.17. Overview of the NO/cGMP signalling pathway.

Despite their lack of protein sequence homology, Cytcp and sGC have a number of similarities, including their ability to bind NO and CO, whilst excluding O\textsubscript{2} from the haem binding site (178, 198). Furthermore, they both exist as high spin ferric and ferrous 5-coordinate species, but do not coordinate H\textsubscript{2}O at the vacant 6\textsuperscript{th} axial position – an unusual feature compared with other haemoproteins. Early UV-vis spectroscopic analyses determined that both sGC and Cytcp haems bind NO with 5-coordinate geometry and CO with a 6-coordinate geometry, which further highlights their similarity in ligand recognition (178, 199-202). Unlike sGC, Cytcp readily crystallises, which allowed the first 3D structure of the protein to be solved by X-ray crystallography in 1980 (203, 204). This Cytcp was isolated from the purple synthetic bacterium \textit{Rhodospirillum molischiaum}, and as predicted from earlier amino acid sequencing and resonance Raman studies, contained a c-type haem, covalently linked to the protein \textit{via} the conserved CxxCH motif, and an axial 5-coordinate histidine ligand bound on the solvent-exposed proximal face of the haem (205, 206). This structure differed significantly from mitochondrial cytochrome \textit{c} in overall fold, sequential location of the CxxCH motif, spin state, and haem solvent exposure (207). Furthermore, although globins also bind uncharged exogenous ligands due to the high spin state haem, they share almost no common features with Cytcp.

Following this initial structural determination, a number of Cytcp isoforms were crystallised from various bacterial strains, which exhibited similar overall fold to one another (208-210). A subsequent crystallographic study of Cytcp isolated from \textit{Alcaligenes xylosoxidans} (AxCytcp), the denitrifying bacterium found in soil, revealed the first structure of NO and CO bound to the haem (211). These structures demonstrated that Cytcp binds CO on the distal face with 6-coordinate geometry, but bound NO on the \textit{proximal} face with 5-
coordinate geometry (211). The fact that NO bound with 5-coordinate geometry was predicted by earlier spectroscopic studies (200), however, the proximal binding and histidine replacement was highly unexpected. A schematic of the ligand bound haem pocket structures are represented in Figure 1.18. Here, haem pocket residues of importance to binding are illustrated, which include solvent-exposed proximal cysteine residues covalently linked to the haem, the 5-coordinate histidine ligand and arginine residue. Whilst on the distal face, a leucine residue facilitates a hydrophobic environment at the vacant 6th coordination site. For CO binding, the leucine residue is displaced, allowing distal binding of CO with an almost linear conformation. For NO binding, the leucine residue is unaffected from the reduced structure, with the proximal histidine displaced by a 5-coordinate NO bound in a bent conformation and the proximal arginine stacked against the haem plane.

Figure 1.18. Haem structure of resting state, CO-, and NO-bound AxCytcp. The figure is based on (211), with the hydrophobic leucine residue shown in red, proximal arginine and cysteine residues shown in green and axial ligands histidine, CO and NO shown in orange.

These findings altered the thinking of the NO binding mechanism, particularly with reference to sGC, where the conformational changes resulting in catalytic production of cGMP could be directly influenced by the displaced histidine residue (192). Therefore, subsequent research focussed on AxCytcp, with a view to elucidating the mechanistic relevance of proximal NO binding. It was illustrated by SF/FT-IR of AxCytcp mixed with NO or $^{15}$NO, that NO binds via a 6-coordinate intermediate (212). These findings were complemented by stopped-flow UV-vis studies, which also identified a 6-coordinate intermediate (213). This resulted in the hypothesis that initially, NO binds on the distal face, which is probed during stopped-flow spectroscopy, followed by displacement of the proximal histidine by a second
NO, stimulating the release of the distally bound NO molecule (214). A schematic of the NO binding mechanism and accompanying description is provided in Figure 5.2. Interestingly, it has also been shown by spectroscopy that sGC also proceeds via a 6-coordinate intermediate during NO binding, further adding to the parallels between sGC and Cytcp (215, 216). The distal hydrophobic leucine residue is thought to be critical during the binding mechanism, as shown from a series of mutagenesis experiments where leucine was replaced with glycine (L16G) and alanine (L16A) (217, 218). During NO binding in the absence of this hydrophobicity, NO is shown to bind distally with a 6-coordinate geometry, apparently unable to disrupt the Fe-His bond sufficiently (218). These studies also illustrated that, unlike the wild-type, the L16A variant is able to bind O₂ at the distal face, suggesting this residue is directly responsible for exclusion of O₂ from the haem pocket (218). Meanwhile, studies of L16 variants with CO returned the highest reported affinity for the ligand by any haemoprotein, with a 10⁸ increase in affinity over the wild-type (217). L16A and L16G variants were isolated with CO bound during overexpression and purification as a direct result of this increase in affinity. Hence, this leucine residue is evidently important in not only the NO binding mechanism, but also the remarkable ligand discrimination demonstrated by this protein.

The basic proximal arginine (R124) was initially suggested to influence Cytcp NO binding due to its rotation upon binding of NO, in doing so assuming a conformation similar to the Cytcp ferric structure (211). A host of variants were generated for this residue in an attempt to identify its role, with the positive charge of arginine converted to: a negative charge with a glutamate residue (R124E); hydrophobic with a phenylalanine (R124F); polar with lysine and glutamine residues (R124K and R124Q, respectively); and indifferent chemistry with an alanine residue (R124A) (219). UV-vis stopped-flow measurements of each of these variants revealed that with greater deviation from the physiological chemistry of the arginine residue, the less pronounced the 6-coordinate intermediate signal (K > Q > F > A > E) (220). Thus, it was hypothesised that arginine, for reasons unknown, stabilises and prolongs the lifetime of the 6-coordinate intermediate during NO binding. Crystallographic studies of each of these variants without NO bound illustrated a highly conserved structure to the wild-type and therefore arginine has no influence in the reduced form (220). This is also the case for the NO-bound form, with the exception of R124A, which returned a mixture of crystal structures. 70 % of R124A structure was in the 5c-NO-bound form on the proximal face as in the wild-type, however, 30 % bound in a 5c-NO conformation on the distal face, with the haem distorted into the proximal cavity vacated by the arginine residue. These structural findings therefore implicate a role for the arginine residue during NO binding of stabilising the haem during histidine bond cleavage and formation of the proximally bound NO species.
The R124 and L16 residues of Cytcp clearly have a significant influence during NO binding. In order to increase their understanding further, the photolysis methods available to us (TA, TRIR and flash photolysis) allows the opportunity to gauge their response following removal of the NO-Fe bond. The first reported photolysis experiment of Cytcp spans back to 1973, when flash photolysis measurements of the Cytcp-CO complex with 100 µs time resolution illustrated that CO is caged in the distal pocket due to a high proportion of prompt recombination with the haem (221). Over the subsequent years, more sophisticated photolysis experiments with the physiologically-implicated NO ligand and wild-type AxCytcp have revealed insights into haem pocket reactivity (222-224). Flash photolysis measurements illustrated that a small percentage of NO escapes to solvent, with the vast majority undergoing geminate recombination, with the histidine recombining in place of any escaped NO via a “kinetic trap” mechanism (222). TA measurements of the AxCytcp-NO complex probed the geminate recombination process directly, confirming that indeed the majority of NO undergoes geminate recombination on a ps timescale (223). In these studies, the role of the arginine residue was suggested to shield against NO solvent escape, although this hypothesis was inconclusive.

The L16A and R124A variants offer the opportunity to probe the influence of these haem binding pocket residues on haem reactivity in the NO complex. By TA, the geminate recombination rates can be compared to those of wild-type previously reported (223), without the potential shielding of arginine in R124A and in the 6c-NO distally bound conformation for L16A. TRIR provides a direct probe of the Fe-NO bond in each case, which can be directly compared to the TA measurements, in addition to providing information on any protein-ligand contacts following bond photolysis. Finally, flash photolysis can be used to determine the increase or decrease of solvent escape in the case of R124A and L16A.
1.4 Thesis Aims and Objectives

UV-vis and IR spectroscopic techniques spanning a wide range of timescales offers the opportunity to increase the mechanistic understanding of proteins containing organometallic chromophores such as tetrapyrrole-based proteins. By studying these proteins using UV-vis spectroscopy, the tetrapyrrole can be directly probed, thus allowing insights into the reaction chemistry by virtue of changes in electronic structure. Whereas IR spectroscopy can be used to directly probe protein dynamics associated with the reaction chemistry; an essential component in facilitating protein function. We therefore aim to test the hypothesis that the combination of these UV-vis and IR time-resolved techniques can help uncover the outstanding mechanistic question related to two tetrapyrrole-based proteins: EAL and cytochrome c'.

Since SF/FT-IR is a relatively novel technique, and this PhD project involved collaboration with the industrial provider of the technology (TgK Scientific), Chapter 2 contains an application note outlining the instrument capabilities and limitations. This characterisation of the technique was important prior to its application to protein systems (particularly EAL), in order to ensure reagent economy and spectroscopic understanding. Chapter 3 focuses on the TRIR and computational analysis of various cobalamin molecules free in solution and acts as a complementary study to the data presented in Chapter 4 involving the TRIR and SF/FT-IR analysis of EAL. The NO-binding protein cytochrome c' is the focus of Chapter 5, and illustrates the potential for probing ligand-haem interactions by photolytic techniques. This chapter also furthers the understanding of Cytcp, which shares mechanistic detail with the therapeutically-implicated NO-sensory protein, sGC. The overall outcome of these results is reviewed in Chapter 6, in addition to providing a view to future research.
2. Stopped-flow FT-IR Application Note: Temporal Resolution and Calibration of a Stopped-flow FT-IR Spectrometer using the Ester Hydrolysis of Methyl Chloroacetate

2.1 Introduction

Infrared (IR) is a vibrational spectroscopy technique that allows the structural analysis of molecules in solution by reporting on the absorbance of diatomic bonds. By coupling a stopped-flow device to a FT-IR spectrometer this allows the experimenter to monitor the formation and cleavage of chemical bonds as they occur. This TgK Scientific/Bruker instrument consists of a stopped-flow sample handling unit connected to a VERTEX 80 FT-IR spectrometer via a thermostatted umbilical supply tube. An integrated mixing cell is mounted in the sample compartment of the spectrometer and consists of CaF$_2$ windows with 100 µm pathlength. Here, the capability of the instrument to monitor IR spectral changes is demonstrated using an ester hydrolysis test reaction. In addition, the effect of spectral resolution and scanner velocity on the spectral acquisition time is covered, along with identification of the optimal shot volume for experimentation.

2.2 Methodology

Hydrolysis of esters has been shown to elicit IR spectral changes in the 1300-1800 cm$^{-1}$ window due to the loss of an ester group and formation of a deprotonated acid (225). The hydrolysis of methyl chloroacetate (MCA) was chosen due to its favourable kinetics (226) for demonstrating the temporal resolution of the SF/FT-IR instrument under pseudo-first order conditions.

2.3 Experimental

MCA was purchased from Sigma-Aldrich (catalogue # 108413) and adjusted to 10 mM with D$_2$O. For hydrolysis, an excess concentration (100 mM) of NaOH was prepared in D$_2$O and all experiments were performed at 25 °C. IR difference spectra (relative to an early time point) for this ester hydrolysis reaction is illustrated in Figure 2.1 measured using 4 cm$^{-1}$ spectral resolution. These difference spectra were generated by averaging 15 individual acquisitions, which reduces the signal to noise ratio. With a single acquisition the noise was ~ 2 mAu, which following averaging was reduced to ~ 0.5 mAu. Averaging of measurements is

---

1This chapter was published as:
possible due to the synchronisation of shot to spectral acquisition in the SF/FT-IR instrument.

Kinetics can be calculated from SF/FT-IR datasets as illustrated in Figure 2.2, which shows plots of absorbance change at distinct IR frequencies. These have been fitted to a single exponential function using a shared rate value, which returned $k_1 = 3.88 \pm 0.03 \text{ s}^{-1}$.

**Figure 2.1.** Average SF/FT-IR difference spectra of MCA vs. NaOH relative to an early time point. These difference spectra were measured at 4 cm$^{-1}$ spectral resolution and are average spectra generated from 15 replicates.

**Figure 2.2.** MCA vs NaOH SF/FT-IR single frequency decays. These traces are generated from the difference spectra illustrated in Figure 2.1, and have been simultaneously fitted to a single exponential function.
The spectral resolution directly influences the number of spectra acquired per second. This correlation is detailed in Table 2.1, and an example of how spectral resolution can affect data acquired is shown in Figure 2.3. This shows the 1599 cm\(^{-1}\) signal decay with various spectral resolutions, and shows with increased spectral quality the number of data points is reduced.

As part of the KinetaDrive software function, the shot volume is variable. In order to determine the optimal shot volume for experimentation, ester hydrolysis was monitored at constant spectral resolution (8 cm\(^{-1}\)) and flow rate (12 ml s\(^{-1}\)), but variable shot volume. Five replicates were acquired for each shot volume and the signal change at 1599 cm\(^{-1}\) analysed by fitting to a single exponential. To estimate data loss, the returned amplitude of signal change was compared with shot volume. An illustration of how the shot volume affects the signal amplitude is illustrated in Figure 2.4, and the plot of amplitude against shot volume in Figure 2.5. This suggests that a shot volume of 100 µl should be used to ensure optimal use of sample whilst providing sufficient volume for an effective displacement of spent solution.

**Table 2.1. Relationship between SF/FT-IR spectral resolution and acquisition rate.**

<table>
<thead>
<tr>
<th>Spectral Resolution</th>
<th>Spectral Acquisition Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 cm(^{-1})</td>
<td>5 spectra/s</td>
</tr>
<tr>
<td>2 cm(^{-1})</td>
<td>9 spectra/s</td>
</tr>
<tr>
<td>4 cm(^{-1})</td>
<td>16 spectra/s</td>
</tr>
<tr>
<td>8 cm(^{-1})</td>
<td>22 spectra/s</td>
</tr>
<tr>
<td>12 cm(^{-1})</td>
<td>26 spectra/s</td>
</tr>
<tr>
<td>16 cm(^{-1})</td>
<td>29 spectra/s</td>
</tr>
</tbody>
</table>

*Figure 2.3. MCA vs. NaOH 1599 cm\(^{-1}\) SF/FT-IR signal measured using various spectral resolutions.*

Each of the traces are fitted to a single exponential. As the spectral resolution is increased there is a loss of kinetic detail with this particular chemical reaction.
Figure 2.4. MCA vs. NaOH 1599 cm\(^{-1}\) SF/FT-IR signal measured using various shot volumes. At low shot volumes there is a loss of signal amplitude at this wavenumber demonstrating a failure to remove spent solution from the transmission cell.

Figure 2.5. The relationship between SF/FT-IR shot volume and signal amplitude. This plot has been constructed using the same data illustrated in Figure 2.4 and demonstrates that for effective measurement of reactions using the SF/FT-IR instrument a shot volume of 100 µl is recommended. The relationship has been fit using a rectangular hyperbola function.

2.4 Results and Conclusion

These data illustrate the capability for the SF/FT-IR instrument to acquire high quality time-resolved IR spectra. The relationship between acquisition and signal quality has been reported, which emphasises the need for a balance between the two. Finally, the optimum shot volume for solvent economy has been estimated at 100 µl.
3. Ultrafast Infrared Spectral Fingerprints of Vitamin B\textsubscript{12} and Related Cobalamins\textsuperscript{2}

3.1 Abstract

Vitamin B\textsubscript{12} (cyanocobalamin, CNCbl) and its derivatives are structurally complex and functionally diverse biomolecules. The excited state and radical pair reaction dynamics that follow their photoexcitation have been previously studied in detail using UV–visible techniques. Similar time-resolved infrared (TRIR) data are limited, however. Herein we present TRIR difference spectra in the 1300–1700 cm\textsuperscript{-1} region between 2 ps and 2 ns for adenosylcobalamin (AdoCbl), methylcobalamin (MeCbl), CNCbl, and hydroxocobalamin (OHCbl). The spectral profiles of all four cobalamins are complex, with broad similarities that suggest the vibrational excited states are related, but with a number of identifiable variations.

The majority of the signals from AdoCbl and MeCbl decay with kinetics similar to those reported in the literature from UV–visible (UV-vis) studies. However, there are regions of rapid (<10 ps) vibrational relaxation (peak shifts to higher frequencies from 1551, 1442, and 1337 cm\textsuperscript{-1}) that are more pronounced in AdoCbl than in MeCbl. The AdoCbl data also exhibit more substantial changes in the amide I region and a number of more gradual peak shifts elsewhere (e.g., from 1549 to 1563 cm\textsuperscript{-1}), which are not apparent in the MeCbl data. We attribute these differences to interactions between the bulky adenosyl and the corrin ring after photoexcitation and during radical pair recombination, respectively. Although spectrally similar to the initial excited state, the long-lived metal-to-ligand charge transfer state of MeCbl is clearly resolved in the kinetic analysis. The excited states of CNCbl and OHCbl relax to the ground state within 40 ps with few significant peak shifts, suggesting little or no homolysis of the bond between the Co and the upper axial ligand. Difference spectra from density functional theory calculations (where spectra from simplified cobalamins with an upper axial methyl were subtracted from those without) show qualitative agreement with the experimental data. They imply the excited state intermediates in the TRIR difference spectra resemble the dissociated states vibrationally (the cobalamin with the upper axial ligand missing) relative to the ground state with a methyl in this position. They also indicate that most of the TRIR signals arise from vibrations involving some degree of motion in the corrin ring. Such coupling of motions throughout the ring makes specific peak assignments neither trivial nor always meaningful, suggesting our data should be regarded as IR spectral fingerprints.

\textsuperscript{2}This chapter was published as:
3.2 Introduction

Vitamin B\textsubscript{12} is one of the most structurally complex vitamins in Nature (Figure 3.1) \cite{227,228}. The central cobalt is coordinated by a corrin ring equatorially, and by a lower-axial 5,6-dimethylbenzimidazole that forms a pseudoribonuclease loop with the corrin. At the ring’s periphery, several acetamide and propionamide side groups project upward and downward, respectively. Although first isolated with an upper axial CN group (i.e., cyanocobalamin, CNCbl) \cite{81}, this form has no known physiological function and it was later discovered that the upper axial position is variable (Figure 3.1) \cite{83}. Adenosylcobalamin (AdoCbl) is the active cofactor in a range of mutase and eliminase enzymes that mediate radical rearrangement reactions. Catalysis by these enzymes is initiated upon substrate binding by homolytic cleavage of the unusual Co–C covalent bond, resulting in a singlet-born, 5′-deoxyadenosyl / cob(II)alamine radical pair. Methylcobalamin (MeCbl), on the other hand, is the cofactor for various methyltransferase enzymes, where the Co–C bond is heterolytically cleaved, thus transferring a methyl cation to the substrate \cite{89,98,109,132,133}. In the absence of protein partners, free AdoCbl is also thought to bind to riboswitch (e.g. \cite{229,230}) regions of mRNA capable of modulating gene expression in response to interaction with metabolites (as opposed to proteins) \cite{231}.

The Co–C bond in AdoCbl and MeCbl can also be homolytically cleaved by light. Aerobic photolysis results in the efficient reaction between the Co\textsuperscript{II} radical and molecular oxygen to ultimately give hydroxocobalamin (OHcbl) \cite{115}. In the absence of oxygen, numerous studies by UV–vis spectroscopy in various time domains (femtosecond pump–probe to continuous wave) have shown that the post-photolysis reaction dynamics of the radical pair involve competition between geminate recombination and cage escape \cite{117-121,126,130,131,156,157,161,232}. Magnetic field effects indicate that, after cage escape, a proportion of separated radical pairs in both AdoCbl and MeCbl also recombine geminately \cite{158,159,163,233}. Such light-sensitivity was initially considered a useful curiosity (a mechanistic tool for detailing the radical pair reaction dynamics) rather than a feature serving an identifiable biological function. However, recent reports suggest that gene regulation may be mediated by a putative AdoCbl-based photoreceptor in the bacteria \textit{Myxococcus xanthus} \cite{122} and \textit{Thermus thermophilus} \cite{124}.

Vibrational spectroscopy has strong potential to help elucidate the exact nature of the interactions between such a structurally complex and functionally diverse cofactor and its dependent enzymes, proteins, and riboswitches. A number of structural markers have been previously identified for vitamin B\textsubscript{12} and its derivatives, including some general assignments for modes from the corrin ring and the peripheral amide groups using FT-IR spectroscopy in the 1500–1700 region \cite{234}. The majority of such work, however, has employed Raman
spectroscopy to focus on the Co–C bond and factors that might influence its lability (e.g. (235-241)). Although the bond was found to weaken with increasing alkyl ligand size, it was relatively unaffected when bound to the enzymes methylmalonyl-CoA mutase (239, 240) and glutamate mutase, (241) or by substitution of the 5,6-dimethylbenzimidazole. Biological chromophores have been studied by ultrafast, time-resolved infrared spectroscopy (TRIR, e.g., free and enzyme bound flavin adenine dinucleotide) (242, 243), but few such investigations exist for cobalamins. The one study of which we are aware was with CNCbl and azidocobalamin, and concentrated on the CN stretching band in the 2050–2200 cm\(^{-1}\) region and the N3 asymmetric stretch in the 1950–2100 cm\(^{-1}\) region, respectively (130). These data suggest the Co–C bond is not appreciably broken in either case but is instead weakened by stretching. We believe there is greater scope for assessing the vibrational changes in vitamin B\(_{12}\) and its derivatives in real time by TRIR: the excited state dynamics, as the Co–C bond is broken and during the ensuing radical pair dynamics. We have therefore probed the IR spectral region between 1300 and 1700 cm\(^{-1}\) up to 2 ns following photoexcitation of AdoCbl, MeCbl, OHClb, and CNCbl, which has provided rich vibrational information about the corrin ring, its amides and the influence of the upper axial ligand. Such vibrational "fingerprints" have the potential to be assessed when the biologically active cofactors, AdoCbl, and MeCbl, are in the presence of their interaction partners, to map any vibrational coupling to biological function.

![Figure 3.1. The chemical structure of cobalamin. Illustrated are the four variable upper axial ligands (R): adenosylcobalamin (coenzyme B\(_{12}\)), cyanocobalamin (vitamin B\(_{12}\)), hydroxocobalamin, and methylcobalamin.](image)
3.3 Materials and Methods

**Materials.** AdoCbl, MeCbl, CNCbl, OHChbl, adenosine, and cis-stilbene were all purchased from Sigma and used without further purification. D$_2$O was purchased from Goss. All samples were prepared in D$_2$O 20 mM HEPES buffer adjusted to pH 7.91 (pD 7.5) with NaOD.

**Static IR Measurements.** Infrared spectra were measured for each of the cobalamin samples ($\sim$8 mM) using a Bruker Vertex 80 spectrometer with an XSA external sample compartment. A CaF$_2$ transmission cell with a 100 μm pathlength was used, which allowed the acquisition of signals in the mid-IR region between 700–4000 cm$^{-1}$.

**TRIR Measurements.** TRIR measurements were performed at room temperature using an experimental setup described previously (64). Amplified Ti:sapphire laser pulses were used to drive optical parametric amplifiers, providing tunable visible and IR output pulses. All cobalamin samples analysed were prepared to a concentration of $\sim$8 mM. The 525 nm excitation pulse energy varied between samples, with 1 μJ used for AdoCbl and MeCbl samples and 500 nJ for OHCbl and CNCbl. Control experiments illustrated that when the pulse energy was 1 μJ or below this did not affect infrared profiles, only the net signal amplitude, with no evidence of multiphoton excitation. For each experiment, the excitation beam was approximately 150 μm fwhm and set at the magic angle with respect to the IR probe beam. The sample cell used had CaF$_2$ windows and a 100 μm path length, and was placed in a rastering sample holder to avoid photobleaching. Difference spectra were measured at time delays after photoexcitation ranging from 100 fs to 2 ns. Changes in IR absorption were measured between 1300 and 1700 cm$^{-1}$ by two overlapping 128 pixel detectors with a resolution of $\sim$3 cm$^{-1}$ per pixel. The pixel–wavenumber calibration was achieved by comparison to the ground state FT-IR spectrum of cis-stilbene.

**Ultrafast UV-vis Measurements.** The ultrafast transient absorption (TA) experiments were performed at room temperature using a setup described previously (163). Briefly, a Ti:sapphire amplifier is pumped by a Q-switched Nd:YLF laser (Positive Light Evolution-30) and seeded by a Ti:sapphire laser (Spectra-Physics Mai Tai). The amplifier has an output of 800 nm, a 1 kHz repetition rate and a 120 fs pulse duration. This beam is then split, with part of the output used to produce tunable radiation in the range 250 to 1000 nm via a noncollinear optical parametric amplifier (Light Conversion TOPAS-White). The other fraction of the amplifier output is used to pump a Helios (Ultrafast Systems LLC) broad-band pump–probe transient absorption spectrometer, with an instrument response function of around 170 fs. For these analyses an excitation wavelength of either 375 or 525 nm was used and absorption changes monitored between 400–700 nm at time delays between 100 fs and
2 ns after excitation. Samples were adjusted to a concentration such that the absorbance at the pump wavelength was ~1.0 using a 2 mm path length quartz cuvette containing a magnetic stirrer bar to avoid photobleaching.

**Computational Chemistry.** Density functional theory (DFT) calculations were performed using either (U)B3LYP/6-31G(d) or (U)BP86/6-31G(d) with Gaussian 03 rev. D.(244) Computed frequencies were scaled, according to literature values (245) by 0.9613 (B3LYP) or 0.9914 (BP86). The presented spectra were generated as the sum of Lorentzian functions with 10 cm$^{-1}$ fwhm.

### 3.4 Results and Discussion

The ground state FT-IR spectra of AdoCbl, MeCbl, OHCbl, and CNCbl in buffered D$_2$O are shown in Figure 3.2, with each normalized to their respective maximum absorbance values. The broad feature in the “fingerprint region” between 1300 and 1500 cm$^{-1}$ is assigned to coupled vibrations in the corrin ring (see DFT calculations). However, the 1550–1700 cm$^{-1}$ window has fewer overlapping features and the large peak at ~1630 cm$^{-1}$ is assigned to amide vibrations of the peripheral acetamide and propionamide groups of the corrin ring. This is consistent with assignments made previously, which also assigned the peak at ~1575 cm$^{-1}$ to breathing modes of the corrin ring (234). The red-shift of the peak at 1624 cm$^{-1}$ for AdoCbl compared with the other cobalamins, is a result of additional signal arising from the adenosyl ligand (Supplementary Information; Figure S3.1). The corrin breathing mode is also red shifted for AdoCbl and MeCbl with respect to CNCbl and OHCbl.

![Figure 3.2. Ground state FT-IR spectra of AdoCbl, MeCbl, OHCbl, and CNCbl. These spectra are normalised to their respective maximum absorbance values, and exhibit similar overall; however, the red-shifted 1624 cm$^{-1}$ AdoCbl peak is likely owing to the axial adenosyl (Figure S3.1). All measurements in this study were performed at room temperature in D$_2$O buffered with 20 mM HEPES, pD 7.5.](image-url)
We will now discuss the TRIR data and kinetic analyses for each cobalamin with reference to published TA spectroscopy data and kinetic parameters. In the case of CNCbl and OHCbl, new TA data will also be presented from measurements made in buffered D$_2$O.

### 3.4.1 Adenosylcobalamin

The TA spectral evolution that follows photolysis of free AdoCbl has been thoroughly analysed by Sension and co-workers (117-121), and Figure 3.3 illustrates an adaptation of their proposed reaction scheme (119). Their analyses show that excitation of ground state AdoCbl is followed by subpicosecond internal conversion to the first excited electronic state, {AdoCbl}*. This then proceeds via a {AdoCbl}** species with calculated rate constants of $k_{ex1} = 0.69 \pm 0.05$ ps$^{-1}$ and $k_{ex2} = 0.071 \pm 0.001$ ps$^{-1}$ to form a {AdoCbl}*** intermediate state within 14 ps in H$_2$O (118). {AdoCbl}*** has a cob(II)alamin-like UV−vis spectral profile, which could represent an excited state with a weakened Co-C bond. This species exists up to 110 ps after excitation, following which it undergoes either relaxation to the ground state (with rate $k_R'$) or homolysis of the weakened Co-C bond resulting in formation of a geminate, singlet-born radical pair ($k_H$). The observed rate of $k_3 = 9.0 \pm 1.0$ ns$^{-1}$ in H$_2$O represents both $k_R$ and $k_H$. It has been previously shown that the fate of this newly formed radical pair is solvent dependent, with greater viscosity increasing the likelihood of geminate recombination (119). The experiments in ref. (119) used solutions of various ethylene glycol content and showed that the rate of cage escape ($k_6$) to form solvent-separated radical pairs decreased from $0.57 \pm 0.06$ ns$^{-1}$ in buffered water to $0.11 \pm 0.03$ ns$^{-1}$ in 100 % ethylene glycol, and the quantum yield of solvent-separated radical pairs fell accordingly from $0.284 \pm 0.020$ to $0.077 \pm 0.020$. The calculated rate constant $k_4 (2.0 \pm 0.2$ ns$^{-1})$ (119) represents both $k_E$ and $k_R$, the latter reporting on the rate of close pair recombination, which is viscosity independent. This reaction mechanism and the associated kinetics are thought to be independent of excitation wavelength between 400 and 530 nm (118). We will use this model (Figure 3.3) as the basis for the analysis of the TRIR data below.
**Figure 3.3. Proposed model for the photolysis of AdoCbl.** This excitation mechanism is adapted from ref (119), with \{AdoCbl\}*, \{AdoCbl\}** and \{AdoCbl\}*** representing the first, second and third excited states of AdoCbl, respectively, each of which has the Co-C bond intact. \(\text{Ado}^* + \text{Cbl}^\|\) represents the AdoCbl singlet-born close radical pair and \(\text{Ado}^* +//+ \text{Cbl}^\|\) the solvent separated radical pair species.

Figure 3.4a shows the ground state FT-IR spectrum of AdoCbl between 1300 and 1700 cm\(^{-1}\), alongside the TRIR difference spectra (with the reference to the ground state) over the 2 ns following excitation at 525 nm (Figure 3.4b and Figure S3.2a). The initial signals (from 2 ps) are those with the largest amplitudes and represent excited state species, which then decay toward the dark state. The spectra are dominated by a bleaching of the band assigned to the corrin breathing mode at around 1575 cm\(^{-1}\) (234), with an accompanying transient at 1549 cm\(^{-1}\), and a bleach and transient at 1500 and 1470 cm\(^{-1}\), respectively. There are peaks that shift gradually over the acquisition period: from 1337 to 1343 cm\(^{-1}\); a peak distortion around 1485 cm\(^{-1}\); a shift from 1549 to 1563 cm\(^{-1}\). The last peak shift is well illustrated when absorbance changes are plotted as a function of time at these two wavenumbers (Figure S3.2b). There are signals at 1442 and 1551 cm\(^{-1}\) that initially decay more quickly (<10 ps), shifting to higher frequencies as they do so, which we attribute to vibrational relaxation in the corrin ring. There is a very rapid peak shift in the opposite direction (to a lower frequency, from 1643 to 1628 cm\(^{-1}\)) that corresponds with the amide I region. The TRIR difference spectra were analysed using single value decomposition (SVD), which identified two principal components (at ~57 and 430 ps) above noise. These basis spectra are illustrated in Figure 3.4c and highlight both the rapid and more gradual peak shifts. On the basis of the time range over which the slower peak shifts occur, they could be reporting on the recombination of the close radical pair to form ground state AdoCbl.
Table 3.1. *AdoCbl TRIR and TA SVD principal kinetics.* Detailed are the calculated rate constants after global fitting of the principal kinetics from SVD analysis of the *AdoCbl TRIR difference spectra*, compared with TA rates reported in (118).^3^

<table>
<thead>
<tr>
<th></th>
<th>TRIR</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$ ($k_{ex1}$)</td>
<td>$1.73 \pm 0.225$ ps$^{-1}$</td>
<td>$0.69 \pm 0.05$ ps$^{-1}$</td>
</tr>
<tr>
<td>$k_2$ ($k_{ex2}$)</td>
<td>$0.229 \pm 0.012$ ps$^{-1}$</td>
<td>$0.071 \pm 0.001$ ps$^{-1}$</td>
</tr>
<tr>
<td>$k_3$ ($k_R + k_{R'\prime}$)</td>
<td>$17.5 \pm 1.54$ ns$^{-1}$</td>
<td>$9.0 \pm 1.0$ ns$^{-1}$</td>
</tr>
<tr>
<td>$k_4$ ($k_R + k_E$)</td>
<td>$2.32 \pm 0.19$ ns$^{-1}$</td>
<td>$2.0 \pm 0.2$ ns$^{-1}$</td>
</tr>
</tbody>
</table>

Figure 3.4. *AdoCbl static FT-IR spectrum; TRIR difference spectra; and SVD principal spectra.* (a) Static FT-IR spectrum of *AdoCbl*. (b) TRIR difference spectra of *AdoCbl* relative to the ground state from 2 ps to 2 ns. Difference spectra have been chosen to illustrate the various peak shifts over time. Fewer spectra with a legend indicating the time delay of each are in Figure S3.2a. (c) Basis spectra corresponding to 57 ps (black) and 430 ps (blue) derived from SVD analysis of the TRIR difference spectra. The scale of the amplitude is arbitrary and does not reflect the relative absorbance of each species; however, the dashed line is at zero, indicating phase.

^3 Refer to Figure 3.3 for the mechanism.
It should be noted that a spectral artefact was observed at 0.8 ps at several wavenumbers (Figure S3.3), which was attributed to cross phase modulation (246). Data at time points 0.8 ps and earlier were therefore omitted from the SVD kinetic analysis. The principal kinetics derived from the SVD basis spectra were globally fit to the sum of four exponentials (Figure 3.5), and the calculated rate constants are given in Table 3.1. The principal components decay with kinetics similar to that of the single wavenumber traces in Figure S3.2b that illustrate the gradual peak shifts in Figure 3.2a. It is likely, therefore, that the black spectrum and kinetic in Figures 3.4c and 3.5, respectively, represent the evolution of the various excited states, and the blue spectrum and kinetic the generation and geminate recombination of the radical pair, which does not decay to zero within 2 ns (119). The data in Table 3.1 are comparable with the kinetic parameters yielded from a similar SVD analysis of the corresponding optical data from the literature (118). The excited state kinetics ($k_{\text{ex1}}$, $k_{\text{ex2}}$ and $k_3$) are all within an order of magnitude, although those from the TRIR data are consistently slightly faster. The difference in $k_{\text{ex1}}$ may be owing to vibrational relaxation, but all may be explained by the potentially different experimental conditions (room temperature etc.) between the studies. The $k_4$ value (corresponding to radical pair recombination and cage escape) is the same within error as the TA kinetic data in ref. (118).

**Figure 3.5. AdoCbl TRIR SVD principal kinetics.** These traces are derived from SVD basis spectra illustrated in Figure 3.4c: 57 ps (black) and 430 ps (blue), fitted globally to the sum of four exponentials. The scale of the amplitude is arbitrary; however, the dashed line is at zero, indicating phase.
3.4.2 Methylcobalamin

The ultrafast photoresponse of MeCbl has been similarly well characterised by TA spectroscopy and differs from that of AdoCbl (118, 120, 121, 126). The photolysis mechanism can proceed via two reaction channels following formation of the initial excited state, with a proportion undergoing "prompt" Co−C bond homolysis resulting in a radical pair species, and the remainder forming a metal-to-ligand charge transfer state (MLCT state) (126). The latter intermediate is suggested to be cob(III)alamin-like with a very weak axial ligand, on the basis of its UV–vis spectral profile, which has a blue-shifted γ band centred at 340 nm. The branching ratio between homolytic cleavage and formation of the MLCT state is excitation wavelength dependent. When excited at 400 nm, roughly a 1:3 ratio is observed, whereas with 520 or 530 nm excitation only the MLCT state is formed (118). These ratios have also been shown to be independent of both solvent (ethylene glycol) and binding to enzyme (methionine synthase) (121, 126). Figure 3.6 outlines the proposed reaction mechanism following photoexcitation of MeCbl. As an excitation wavelength of 525 nm has been used in all our TRIR experiments, the "prompt" homolysis reaction channel has been omitted, with all the initial excited states converted to the MLCT state, \{H₃C⁻−Cbl\} (126). This species exists for ~1 ns, following which it undergoes a second partition, with 86 % relaxing to the ground state molecule and the remaining 14 % undergoing Co−C bond homolysis. Subsequently, this radical pair exhibits no significant geminate recombination, with the vast majority forming solvent separated radical pairs (126). This is likely to be owing to the small size and highly reactive nature of methyl radicals, resulting in rapid diffusion into a solvent separated state (120). However, it has also been shown by magnetic field effect studies that radical recombination does occur following solvent separation (159). Published kinetic parameters for photolysis of MeCbl were generated by SVD analysis of TA spectroscopy, followed by global fitting to the sum of three exponentials. These yielded rate constants of 0.68 ± 0.30 ps⁻¹, 0.060 ± 0.010 ps⁻¹, and 1.10 ± 0.15 ns⁻¹, the last of which is likely to represent the sum of \(k_R\) and \(k_H\) (118).

The FT-IR ground state and TRIR difference spectra from 2 ps after excitation at 525 nm are shown in Figure 3.7a,b, respectively. Similar signal bleaches and transients to the AdoCbl data dominate the TRIR difference spectra. There are, however, subtle but significant differences. The gradual peak shifts apparent in the AdoCbl spectra are largely absent for MeCbl; for example, the signals at 1337 and 1550 cm⁻¹ relax to the ground state without shifting. On the basis of the current reaction scheme, this would suggest the initial excited state and MLCT intermediate have similar IR signal differences relative to the ground state. Also, the lack of significant geminate recombination from the close pair further supports the assertion that the slower peak shifts evident in the AdoCbl TRIR difference spectra are a
result of interactions between the bulky adenosyl and the corrin ring during recombination. Another variation is around 1640 cm\(^{-1}\), where in the case of AdoCbl significant signal changes were observed, and there is an initial, rapid peak shift to a lower frequency. For MeCbl, the changes are less pronounced, with the initial, fast peak shift being more modest and to a higher frequency (from 1646 to 1649 cm\(^{-1}\)). These signals are therefore more consistent with vibrational relaxation of peripheral amide groups, whereas the larger signal changes observed for AdoCbl in this region may correspond to interactions between the upper axial adenosyl and the corrin amides following excitation. A common spectral feature is the presence of the fast signal relaxation at 1422 cm\(^{-1}\), which decays in <10 ps of excitation in both AdoCbl and MeCbl and can therefore be attributed to fast vibrational relaxation of the corrin ring.

\[ \text{Figure 3.6. Proposed model for the photolysis of MeCbl. This excitation pathway is adapted from (118) and represents photolysis with 525 nm excitation. The “prompt” homolysis partition that is known to manifest after excitation at shorter wavelengths (around 400 nm) has been omitted due to the 525 nm excitation wavelength used here.} \]

\[ \text{Figure 3.7. MeCbl static FT-IR spectrum and TRIR difference spectra. (a) Static FT-IR spectrum of MeCbl. (b) TRIR difference spectra of MeCbl relative to the ground state from 2 ps to 2 ns. Difference spectra have been chosen to illustrate the lack of peak shifts over time compared to AdoCbl (Figure 3.4b). Fewer spectra with a legend indicating the time delay of each are in Figure S3.4.} \]
SVD analysis yielded only one basis spectrum, which has the same profile as the initial difference spectrum. Possible explanations for the lack of more basis spectra include the structural similarity between the initial excited state and MLCT state, the relative lack of radical pair recombination, and the relatively modest signals (in comparison to AdoCbl) representing the faster decay components. The principal kinetic was fit to the sum of three exponentials (Figure 3.8) and the calculated rate constants detailed in Table 3.2. The decay profile has three distinct phases: the signal manifests on a fast time scale (~1 ps) and then decays slightly over ~100 ps before tending toward zero over the following ~2 ns. The intermediate is likely to be the long-lived MLCT intermediate that forms before relaxing back to the ground state or forming solvent-separated radicals (Figure 3.6). This is in agreement with the equivalent TA spectroscopy findings (118, 126); where the rate constants (Table 3.2) correspond within error, apart from \( k_{\text{ex1}} \), which is faster in the TRIR and may represent vibrational relaxation.

**Table 3.2. MeCbl TRIR and TA SVD principal kinetics.** The TRIR rate values were calculated by global fitting of the SVD principal kinetics of the MeCbl TRIR difference spectra. These are then compared with TA rates reported in (118).^4^

<table>
<thead>
<tr>
<th></th>
<th>TRIR</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k_1 ) (( k_{\text{ex1}} ))</td>
<td>1.92 ± 0.16 ps(^{-1})</td>
<td>0.68 ± 0.30 ps(^{-1})</td>
</tr>
<tr>
<td>( k_2 ) (( k_{\text{ex2}} ))</td>
<td>0.0357 ± 0.041 ps(^{-1})</td>
<td>0.060 ± 0.010 ps(^{-1})</td>
</tr>
<tr>
<td>( k_3 ) (( k_{\text{H}} + k_{\text{R'}} ))</td>
<td>1.08 ± 0.20 ns(^{-1})</td>
<td>1.10 ± 0.15 ns(^{-1})</td>
</tr>
</tbody>
</table>

**Figure 3.8. MeCbl TRIR SVD principal kinetic trace.** This decay is plotted on a logarithmic time scale and fitted to the sum of three exponentials. The data point at 1.8 ps has been excluded from the fitted data as the fit failed to minimize with its inclusion. The scale of the amplitude is arbitrary; however, the dashed line is at zero, indicating phase.

^4 Refer to Figure 3.6 for the mechanism.
3.4.3 Cyanocobalamin and Hydroxocobalamin.

There are also previous reports of the ultrafast photoresponses of both CNCbl and OHCbl \((130, 131)\). Although neither acts as functional coenzymes, CNCbl is an essential human dietary supplement, whereas OHCbl is a primary photoproduct of AdoCbl and MeCbl after aerobic photolysis in aqueous solution. The photoresponses of both CNCbl and OHCbl proceed via a single excited state formed on a subpicosecond time scale \((130)\). The identity of the CNCbl excited state has been previously explored using TRIR analyses, which focused on the CN stretch band at 2138 cm\(^{-1}\) \((130)\). Following excitation, this band was red-shifted to 2120 cm\(^{-1}\), a position significantly higher than that of CN\(^{-}\) at 2080 cm\(^{-1}\) and CN radical at 2046 cm\(^{-1}\). On the basis of these findings it was theorised that photoexcitation does not result in Co–C bond cleavage, but instead CNCbl undergoes a fast electronic internal conversion of a corrin ring \(\pi\) to Co 3dz\(^2\) ligand to metal charge transfer (LMCT), resulting in weakening and lengthening of the Co–C bond. This hypothesis is also supported by TD-DFT calculations \((247)\). The CNCbl excited state was shown to be short-lived, with a calculated rate constant of \(0.149 \pm 0.018\) ps\(^{-1}\) in \(H_2O\). Furthermore, the excited state lifetime is independent of excitation wavelength, but dependent upon the solvent dielectric constant, to which it is inversely related \((131)\). On the other hand, a recent study has shown that when plasmid DNA is incubated with OHCbl under continuous illumination this results in DNA cleavage, suggesting OH radical formation \((116)\). However, these analyses required prolonged exposure (up to 1 hour) to high intensity polychromatic light, suggesting a low quantum yield of OH radicals after photoexcitation.

The ground state FT-IR spectra of CNCbl and OHCbl are shown in Figures 3.9a and 3.10a, respectively, and exhibit very high similarity to one another and to that of MeCbl. TRIR data are shown in Figure 3.9b for CNCbl and in Figure 3.10b for OHCbl. Early time spectra show broadly similar peak positions in the fingerprint region to AdoCbl and MeCbl, but with different relative amplitudes, especially for CNCbl. For example, there is still a strong bleach at \(~1575\) cm\(^{-1}\) in the CNCbl data, but the amplitude of the accompanying transient is relatively small by comparison. This is also true (but to a lesser extent) for OHCbl and may be a feature that indicates the likelihood of bond homolysis. The TRIR signals at 1500, 1549, and 1573 cm\(^{-1}\) in AdoCbl and MeCbl are red-shifted in the CNCbl spectra to 1504, 1560, and 1576 cm\(^{-1}\), respectively, and to a similar extent in the OHCbl spectra. Another qualitative difference for CNCbl includes a peak shift during relaxation from 1640 to 1626 cm\(^{-1}\) in the amide region, which is not present in OHCbl. For both CNCbl and OHCbl the excited state is very short-lived, with the signals returning to the ground state within 40 ps.
Figure 3.9. CNCbl static FT-IR spectrum and TRIR difference spectra. (a) FT-IR ground state spectrum of CNCbl. (b) Selected TRIR difference spectra for CNCbl relative to ground state from 0.7 to 40 ps. Fewer spectra with a legend indicating the time delay of each are in Figure S3.5.

Figure 3.10. OHCbl static FT-IR spectrum and TRIR difference spectra. (a) FT-IR ground state spectrum of OHCbl. (b) Selected TRIR difference spectra for OHCbl relative to the ground state from 0.7 to 40 ps. Fewer spectra with a legend indicating the time delay of each are in Figure S3.8.

The TRIR difference spectra were analysed by SVD, which returned one basis spectrum above noise for both CNCbl and OHCbl, with the same overall profile as their respective difference spectra. The principal kinetics were fit to the sum of two exponentials in each case and rate constants calculated. Because the rate constants for relaxation of CNCbl are solvent dielectric constant dependent (130), equivalent TA spectroscopy measurements were taken for CNCbl and OHCbl in D$_2$O at excitation wavelengths 375 and 525 nm to allow direct comparison with the TRIR data (Figures S3.6, S3.7, S3.9, and S3.10). These yielded difference spectra relative to the ground state of similar profile to previous studies (130).
When excited at 525 nm, there was significant scatter from the pump beam between ~510–540 nm, and therefore 375 nm excitation was employed to yield complete difference spectra in the 425–650 nm window. Rate constants extracted at 470, 490, and 553 nm from CNCbl and 455 and 470 nm from OHCbl (Tables S3.1 and S3.2, respectively) from overlapping regions in the 375 and 525 nm excitation data suggest that, in accordance with the literature, decay kinetics are excitation wavelength independent. Therefore, TA data from 375 nm excitation were also analysed using SVD, the principal kinetics of which were fit globally to the sum of two exponentials and the rate constants compared with those from the TRIR data (Table 3.3). The observed values for \( k_2 \) for both CNCbl and OHCbl are the same within error between the infrared and optical measurements, and the values for \( k_1 \) are in the same order. It is therefore reasonable to assume that the observed processes are comparable.

**Table 3.3. CNCbl and OHCbl TRIR and TA SVD principal kinetics.** Samples were prepared in D\(_2\)O buffer and excited at 525 nm for TRIR measurements and 375 nm for the TA.

<table>
<thead>
<tr>
<th></th>
<th>TRIR</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNCbl</td>
<td>( 0.532 \pm 0.059 ) ps(^{-1} )</td>
<td>( 0.129 \pm 0.004 ) ps(^{-1} )</td>
</tr>
<tr>
<td>OHCbl</td>
<td>( 1.14 \pm 0.44 ) ps(^{-1} )</td>
<td>( 0.136 \pm 0.020 ) ps(^{-1} )</td>
</tr>
</tbody>
</table>

3.4.4 Model Frequency Calculations

The reported static and TRIR spectra of the four cobalamins in this study are clearly quite complex, yet all have generally similar spectral properties. Although the broad 1600–1700 cm\(^{-1} \) band is likely to be dominated by amide I resonances from the peripheral acetamide and propionamide groups (234), further peak assignments by analogy are not trivial. As a first step to the assignment of the TRIR spectral features, a series of minimal cobalamin DFT models (Figure 3.11) were examined. These were built by pruning all heavy atoms at a distance of two atoms or more from the corrin ring. Similar models have been used to successfully investigate the Raman spectra of several cobalamins (e.g., (248, 249)). We performed gas phase calculations using both B3LYP/6-31G(d) and BP86/6-31G(d) (Figures 3.12, S3.11 and S3.12). Both functionals were used as BP86 has been shown to perform better in some cases than B3LYP for cobalamin models (250). Electronic excited state calculations (e.g., using TD-DFT) are beyond the scope of the present work. Instead, intermediate species were crudely approximated using Cbl-Co\(^{III} \) and Cbl-Co\(^{II} \) models, which lack the sixth axial ligand as a result of heterolysis and homolysis, respectively. As the low frequency region (i.e., the Co–N and Co–R stretches) is not the focus of the present study, this approach is
reasonable if the Co-R bond of the intermediates is significantly weakened following photoexcitation. Due to the omission of the acetamide and propionamide groups in the model, the 1600−1700 cm$^{-1}$ region of the calculated spectra is expected to be sparse. As the AdoCbl TRIR data are more complex than the other three cases (i.e., there are 2 spectral intermediates observed; Figure 3.4c), we will focus on CH$_3$Cbl, CNCbl, and OHCbl models. No attempt to scale the calculated spectra to fit experiment was made and computed frequencies were scaled according to literature values (245).

Both the radical (Cbl-Co$^{II}$) and positively charged (Cbl-Co$^{III}$) sixth ligand-free species were also examined to account for homolytic and heterolytic cleavage, respectively, of the Co-R bond.

Spectra for the CH$_3$Cbl, Cbl-Co$^{III}$, and Cbl-Co$^{II}$ B3LYP models are shown in Figure 3.12 and Figure S3.11; we found these to be more similar to the experimental spectra than the BP86 calculations (Figure S3.12). Both the Cbl-Co$^{III}$ − CH$_3$Cbl and the Cbl-Co$^{II}$ − CH$_3$Cbl difference spectra in Figure 3.12b show some qualitative agreement with the TRIR data, especially those of AdoCbl and MeCbl (Figures 3.4b and 3.7b, respectively); most prominently, negative peaks at ~1558 and 1500 cm$^{-1}$ and a positive peak at ~1550 cm$^{-1}$. These data would suggest that the various intermediate states, including the initial excited state intermediates, the MLCT state of MeCbl and the radical pair photoproducts, resemble the Cbl-Co$^{III}$ or Cbl-Co$^{II}$ dissociated state models vibrationally. They also suggest that the majority of the TRIR signal arises from vibrations in the corrin ring, rather than its peripheral substituents. The Cbl-Co$^{III}$ − Cbl-Co$^{II}$ difference spectrum in Figure 3.12b, however, deviates more substantially from the TRIR data.

Figure 3.11. Cbl model structure used for DFT calculations. The sixth axial ligand, R = CH$_3$, CN, OH. Spectra for the CH$_3$Cbl, Cbl-Co$^{III}$, and Cbl-Co$^{II}$ B3LYP models are shown in Figure 3.12 and Figure S3.11; we found these to be more similar to the experimental spectra than the BP86 calculations (Figure S3.12). Both the Cbl-Co$^{III}$ − CH$_3$Cbl and the Cbl-Co$^{II}$ − CH$_3$Cbl difference spectra in Figure 3.12b show some qualitative agreement with the TRIR data, especially those of AdoCbl and MeCbl (Figures 3.4b and 3.7b, respectively); most prominently, negative peaks at ~1558 and 1500 cm$^{-1}$ and a positive peak at ~1550 cm$^{-1}$. These data would suggest that the various intermediate states, including the initial excited state intermediates, the MLCT state of MeCbl and the radical pair photoproducts, resemble the Cbl-Co$^{III}$ or Cbl-Co$^{II}$ dissociated state models vibrationally. They also suggest that the majority of the TRIR signal arises from vibrations in the corrin ring, rather than its peripheral substituents. The Cbl-Co$^{III}$ − Cbl-Co$^{II}$ difference spectrum in Figure 3.12b, however, deviates more substantially from the TRIR data.
The calculated CNCbl and OHCbl spectra are compared to the CH$_3$Cbl spectrum in Figure S3.13. Over the spectral window available for the TRIR experiments, the three spectra are very similar. The prominent CH$_3$Cbl peak at 1558 cm$^{-1}$, which is comparable to the $\sim$1575 cm$^{-1}$ band previously assigned to a corrin breathing mode (234), shifts to 1562 and 1566 cm$^{-1}$ in the CNCbl and OHCbl spectra, respectively, which is consistent with the FT-IR data in Figure 3.2. Elsewhere, the CN stretch occurs at 2183 cm$^{-1}$ (cf. the experimental value of 2138 cm$^{-1}$ (130)) and the axial OH at 3596 cm$^{-1}$. The CH$_3$Cbl 1558 cm$^{-1}$ peak red-shifts in the Cbl-Co$^{\text{III}}$ and Cbl-Co$^{\text{II}}$ spectra (Figure 3.12a) generating a similar bleach and transient in Figure 3.12b that is observed in the TRIR data for all four cobalamins. In our calculations, this arises from a mode in the corrin that oscillates in the plane of the ring. The prominent CH$_3$Cbl peaks at 1486 and 1358 cm$^{-1}$ involve multiple CH$_3$ bending modes that are coupled through the ring. As these involve CH$_3$ groups created during the “pruning” process while creating the Cbl model (cf. acetamide and propionamide groups), they may be a model artefact. However, all of the calculated vibrational modes within the 1300–1700 cm$^{-1}$ experimental frequency range (there are 66 in the CH$_3$Cbl model) involve some degree of motion in the corrin ring. Therefore, it seems that the TRIR data in this study should be interpreted as “fingerprints” rather than in the context of specific peak assignments.

Figure 3.12. BP86 DFT calculated spectra for the Cbl model structure. (a) Computed vibrational spectra of the Cbl models (Figure 3.11) in the experimental region: CH$_3$Cbl (black); Cbl-Co$^{\text{II}}$ (red); Cbl-Co$^{\text{III}}$ (blue). (b) Difference spectra: Cbl-Co$^{\text{III}}$ – CH$_3$Cbl (blue); Cbl-Co$^{\text{II}}$ – CH$_3$Cbl (red); Cbl-Co$^{\text{III}}$ – Cbl-Co$^{\text{II}}$ (green). A broader range is represented in Figure S3.11.
3.5 Summary and Conclusions

We have presented static and TRIR spectra of AdoCbl, MeCbl, CNCbl, and OHCbl between 1300 and 1700 cm$^{-1}$. This region encompasses many of the main vibrations in the cobalamin chromophore, including the corrin ring and its amide substituents. Although the spectral profiles of each cobalamin are dominated by comparable features, there are identifiable variations. The bulky adenosyl ligand contributes both directly and sterically to the AdoCbl signals. This is evinced by peak shifts in both its static FT-IR spectrum (1630 to 1624 cm$^{-1}$) relative to the other cobalamins, and its TRIR spectra, where a rapid peak shift to lower frequencies (1643 to 1628 cm$^{-1}$) immediately follows excitation, and a number of gradual peak shifts (e.g., 1549 to 1563 cm$^{-1}$) accompany radical pair recombination. The MLCT metastable excited state that forms following excitation of MeCbl ($126$) is resolved kinetically, despite the fact that its difference spectra relative to the ground state closely resemble those of the initial excited state. In fact, the kinetics of the main decay components of all of the cobalamins studied are comparable to those calculated from the TA data reported by Sension and co-workers, and these vibrational data are therefore consistent with their proposed photophysical models ($126, 130, 131, 156$). The spectral features, however, are numerous, of high quality and likely to hold further detail, not all of which can be fully realised here. For instance, the relative amplitudes of the major bleach at $\sim 1575$ cm$^{-1}$ and its accompanying transient at $\sim 1549$ cm$^{-1}$ appear to relate to the likelihood of homolysis following photoexcitation. Further, there are additional, rapidly decaying signals in the AdoCbl and MeCbl TRIR spectra that we attribute to vibrational relaxation in the corrin ring and amides. Consequently, there is great potential for assessing the vibrational interactions between the biologically active B$_{12}$ species and their various biological partners. The amide groups of the corrin, for instance, are known to be heavily involved in protein binding, but the precise role of the rest of the corrin ring in biological function is yet to be determined. Our DFT calculations reproduce many of the features observed in both the FT-IR and TRIR data. They suggest the excited state intermediates in the TRIR difference spectra resemble the dissociated states (the cobalamin with the upper axial ligand missing) relative to the ground state with a methyl in this position. They also indicate that most of the TRIR signals arise from vibrations involving some degree of motion in the corrin ring. Such coupling of motions throughout the ring makes specific peak assignments neither trivial nor always meaningful, suggesting our data should be regarded as IR spectral fingerprints. Changes to these fingerprints upon binding to a protein or riboswitch, or during catalysis, may help uncover the role of the corrin in biological function. Such changes may also in turn shed more light on the various ring modes themselves, by altering the extent or nature of the coupled motions.
3.6 Supplementary Information

Figure S3.1. Ground state FT-IR spectra for adenosine and AdoCbl. The spectra for adenosine (black trace) and AdoCbl (red trace) have both been normalised to their maximum and minimum absorbance. Adenosine and AdoCbl share the same peak position at 1624 cm\(^{-1}\) and hence it is likely the axial adenosyl group is responsible for the red shift in this region for AdoCbl relative to the other cobalamins studied.

Figure S3.2. AdoCbl TRIR difference spectra and single frequency decays. (a) Selected TRIR difference spectra relative to the ground state from 2 ps – 2 ns following photoexcitation of AdoCbl at 525 nm. (b) Single wavenumber decays from the AdoCbl TRIR difference spectra plotted on a logarithmic time scale. The traces illustrate the peak shift between 1549 cm\(^{-1}\) (transparent diamonds) and 1562 cm\(^{-1}\) (filled circles) in the TRIR difference spectra following excitation of AdoCbl at 525 nm. They are globally fit to the sum of 4 exponentials, with rate constants similar to those derived from SVD.
Figure S3.3. AdoCbl TRIR single wavenumber traces at 1472 and 1551 cm\(^{-1}\). Both these signal decays exhibit a “spike” at 0.8 ps probably resulting from cross phase modulation (246).

Figure S3.4. Selected time point MeCbl TRIR difference spectra. These difference spectra are relative to the ground state from 2 ps – 2 ns following photoexcitation of MeCbl at 525 nm.
Figure S3.5. Selected time point CNCbl TRIR difference spectra. These difference spectra are relative to the ground state from 2 ps – 2 ns following photoexcitation of CNCbl at 525 nm.

Figure S3.6. CNCbl TA difference spectra with 375 nm photoexcitation. These difference spectra are between 1 – 40 ps following photoexcitation of CNCbl using a 375 nm pump wavelength. The early time spectra are those with the largest amplitude.
Table S3.1. **CNCbl 375 and 525 nm excitation TA spectroscopy rate constants at various single wavelengths.** The data at 470 and 490 nm fit to one exponential, whereas the data at 553 nm fit to the sum of 2 exponentials. These positions were chosen as they report signal changes unaffected by scattered light from the pump beam. The rate constants are the same within error.

<table>
<thead>
<tr>
<th></th>
<th>375 nm excitation</th>
<th></th>
<th>525 nm excitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_1$</td>
<td>$k_2$</td>
<td>$k_1$</td>
</tr>
<tr>
<td>470 nm</td>
<td>0.135 ± 0.002 ps$^{-1}$</td>
<td>-</td>
<td>0.136 ± 0.002 ps$^{-1}$</td>
</tr>
<tr>
<td>490 nm</td>
<td>0.130 ± 0.002 ps$^{-1}$</td>
<td>-</td>
<td>0.137 ± 0.002 ps$^{-1}$</td>
</tr>
<tr>
<td>553 nm</td>
<td>0.569 ± 0.095 ps$^{-1}$</td>
<td>0.128 ± 0.003 ps$^{-1}$</td>
<td>0.638 ± 0.015 ps$^{-1}$</td>
</tr>
</tbody>
</table>

**Figure S3.7. CNCbl TA difference spectra with 525 nm photoexcitation.** These difference spectra are between 1 - 40 ps following photoexcitation of CNCbl using a 525 nm pump wavelength. The early time spectra are those with the largest amplitude. The distorted signal between 510-540 nm is due to scatter from the pump beam.
Figure S3.8. **Selected time point OHCbl TRIR difference spectra.** These difference spectra are relative to the ground state from 2 ps – 2 ns following photoexcitation of OHCbl at 525 nm.

Figure S3.9. **OHCbl TA difference spectra with 375 nm photoexcitation.** These difference spectra are between 1 – 40 ps following photoexcitation of OHCbl using a 375 nm pump wavelength. The early time spectra are those with the largest amplitude.
Table S3.2. OHCbl 375 and 525 nm excitation TA spectroscopy rate constants at various single wavelengths. These positions were chosen as they report signal changes unaffected by scattered light from the pump beam. The rate constants are the same within error.

<table>
<thead>
<tr>
<th></th>
<th>375 nm excitation</th>
<th>525 nm excitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_1$</td>
<td>$k_2$</td>
</tr>
<tr>
<td>455 nm</td>
<td>$0.375 \pm 0.112$ ps$^{-1}$</td>
<td>$0.098 \pm 0.021$ ps$^{-1}$</td>
</tr>
<tr>
<td>470 nm</td>
<td>$0.371 \pm 0.073$ ps$^{-1}$</td>
<td>$0.082 \pm 0.013$ ps$^{-1}$</td>
</tr>
</tbody>
</table>

Figure S3.10. OHCbl TA difference spectra with 525 nm photoexcitation. These difference spectra are between 1 – 40 ps following photoexcitation of OHCbl using a 525 nm pump wavelength. The early time spectra are those with the largest amplitude. The distorted signal between 510-540 nm is due to scatter from the pump beam.
Figure S3.11. Broad wavenumber BP86 DFT calculated spectra for the Cbl model structure. These spectra are expanded versions of Figure 3.12, which illustrates additional calculated IR signals outside of the experimental spectral window (1300-1700 cm\(^{-1}\)). (a) Computed vibrational spectra of the Cbl models (Figure 3.11): CH\(_3\)Cbl (black); Cbl-Co\(^{II}\) (red); and Cbl-Co\(^{III}\) (blue). (b) Difference spectra: Cbl-Co\(^{III}\) - CH\(_3\)Cbl (blue); Cbl-Co\(^{II}\) - CH\(_3\)Cbl (red); and Cbl-Co\(^{III}\) - Cbl-Co\(^{II}\) (green).

Figure S3.12. BP86/6-31G(d) DFT calculated spectra for the Cbl model structure. These vibrational spectra represent CH\(_3\)Cbl (black); Cbl-Co\(^{II}\) (red); and Cbl-Co\(^{III}\) (blue) calculated using the BP86/6-31G(d) computer model.
Figure S3.13. B3LYP/6-31G(d) DFT calculated spectra for the Cbl model structure. These vibrational spectra represent CH₃Cbl (black), OHCbl (red) and CNCbl (blue) models calculated using B3LYP/6-31G(d). The peak at 1200 - 1250 cm⁻¹ arises from coupling between the 5th axial ligand and the corrin ring and is, in part, an artefact of using NH₃ (cf. DMB) as the ligand.
4. Protein Motions Are Coupled to the Reaction Chemistry in Coenzyme B\textsubscript{12}-Dependent Ethanolamine Ammonia Lyase\textsuperscript{5}

4.1 Introduction

Coenzyme B\textsubscript{12} (5'-deoxyadenosylcobalamin, AdoCbl)-dependent enzymes facilitate radical rearrangement reactions in a variety of organisms (\textsuperscript{132, 133}). Upon substrate binding, the unique covalent Co-C bond at the centre of AdoCbl is broken homolytically, yielding a singlet-born 5'-deoxyadenosyl and cob(II)alamin radical pair (\textsuperscript{132, 133}). Remarkably, thermal homolysis rates achieved by these enzymes are about $10^{12}$ times greater than that of the free cofactor in solution (\textsuperscript{159, 251-253}). Recent data indirectly suggest protein motions may be coupled to the reaction chemistry (\textsuperscript{163}), a hotly debated issue in modern enzymology (\textsuperscript{254, 255}). Herein, from studies using both laser flash photolysis and stopped-flow infrared (IR) techniques, we present the first direct evidence for a protein motion that correlates with coupled Co-C bond homolysis and H abstraction from the substrate in ethanolamine ammonia lyase (EAL) from \textit{Salmonella enterica}.

EAL is in many ways typical of AdoCbl-dependent enzymes that catalyse isomerization and elimination reactions (\textsuperscript{163}). After Co-C homolysis, the adenosyl radical abstracts hydrogen from ethanolamine, followed by radical rearrangement and dissociation to products (Figure 4.1) (\textsuperscript{98, 154}). In theory, the only steps resolvable by UV-visible spectroscopy are the interconversion of cob(III)alamin and cob(II)alamin at the beginning and end of turnover. However, much like methylmalonyl-CoA mutase (\textsuperscript{251}) and glutamate mutase (\textsuperscript{252}), homolysis of the Co-C bond and H abstraction from the substrate are kinetically coupled in EAL, such that kinetic isotope effects are observable in the signal representing homolysis (\textsuperscript{154, 159}). This coupling has been shown to remove magnetic field-sensitivity from the Co-C homolysis in EAL, by limiting the extent of geminate recombination, and thus favouring the dissociated state (\textsuperscript{158, 159}). EPR spectroscopy has also shown the substrate radical to accumulate during turnover and that it is separated from the Co\textsuperscript{II} by 8.7 Å (\textsuperscript{256}), further stabilising against geminate recombination. Although these factors no doubt contribute to the catalytic power of EAL, there remains little evidence for an explicit role for the protein. Unlike methylmalonyl-CoA mutase (\textsuperscript{96}), diol dehydratase (\textsuperscript{257}) and ornithine 4,5-aminomutase (\textsuperscript{258, 259}), there is yet no evidence of any large-scale motions associated with substrate binding in EAL (\textsuperscript{101}). Instead, Robertson et al. have proposed a more subtle,

\textsuperscript{5}This chapter was published as:
cooperative effect from the protein (162), and, in support of this, we have recently published evidence that implicates a dynamic contribution from an active-site residue (163).

![Diagram of reaction mechanism for EAL]

**Figure 4.1. Current reaction mechanism for EAL.** Adapted from (154). Substrate binding initiates homolysis, with the 5′-deoxyadenosyl radical generating the substrate radical by H abstraction, followed by rearrangement to the product radical, which then dissociates to the acetaldehyde and ammonia.

EAL exhibits three broad amide absorptions in the mid-IR spectrum (Figure 4.1a, solid line) with the amide I band (1625–1675 cm⁻¹) representing the most intense signal. AdoCbl also contributes to the signal in this region (Figure 4.1a, dashed line), with a broad band centred around 1625 cm⁻¹ (likely owing to amide groups of the corrin ring). AdoCbl also absorbs strongly in the fingerprint region (1300–1500 cm⁻¹), where many vibrational modes overlap (260). IR spectroscopy has been used to good effect in AdoCbl-dependent systems previously; however, without time resolution (240, 241). In an attempt to directly assess the influence of protein motions on the reaction chemistry in EAL, the Scrutton group have employed two dynamic IR methods. A stopped-flow sample handling unit coupled to a Fourier transform IR spectrometer (SF/FT-IR) allowed spectral changes to be monitored following rapid mixing of reactants (46, 47). The ultrafast time-resolved IR (TRIR) spectrometer, on the other hand, employs a pump-probe setup and enabled spectral measurements with femtoseconds time resolution after photoexcitation of the sample (64). TRIR has been previously employed to probe cofactor–protein interactions in systems such as bacteriorhodopsin (261, 262). As will become evident, such interactions are subtle during photolysis of EAL-bound AdoCbl. In this case, therefore, the TRIR data serve to illustrate the
distinction between Co-C homolysis by light and by substrate binding. Overall signals in the spectral window measured by TRIR are very different to those measured by SF/FT-IR experiments. We conclude that substrate binding, therefore, has a considerably stronger impact on the protein, which appears to move in conjunction with the reaction chemistry. Comparing TRIR measurements of free and EAL-bound AdoCbl provided the opportunity to determine whether there is interaction between the protein and cofactor following photoexcitation and Co-C bond photolysis.

4.2 Materials and Methods

Materials. E. coli was transformed with the plasmid ‘pET SEAL’ which encodes the small (31 kDa) and large (49 kDa) subunits of EAL from Salmonella enterica. This was then overexpressed and purified as previously described (263). AdoCbl, ethylene glycol, ethanolamine and 2-aminopropanol were purchased from Sigma without further purification. All samples were prepared in D₂O 20 mM HEPES buffer adjusted to pH 7.9 (pD 7.5) with NaOD. For preparation of the EAL holoenzyme sample, the apoenzyme was incubated with an excess of AdoCbl (400 μM) on ice for 10 minutes under safe light. This was then passed down an Econo-pac® DG10 column pre-equilibrated with D₂O HEPES buffer described above to remove any unbound AdoCbl. The eluted EAL sample active site concentration was determined by its 525 nm absorbance reading (ε = 8 mM⁻¹ cm⁻¹). Preparation of ¹⁵N isotopically-labelled EAL used a minimal medium protocol with ¹⁵N-labelled ammonium chloride as the nitrogen source. The transformation protocol was as previously described (263), but all prepared cultures utilised minimal medium to ensure the only source of nitrogen was isotopically labelled. 500 ml of minimal medium was produced by preparation of two solutions: “Solution A” contained 100 mM Na₂HPO₄ and 40 mM KH₂PO₄ in 487.5 ml which was adjusted to pH = 7.2 then autoclaved; “Solution B” contained 160 mgml⁻¹ glucose, 9.6 mgml⁻¹ MgSO₄, 40 mgml⁻¹ ¹⁵NH₄Cl, 0.1 mgml⁻¹ MnSO₄, 0.75 mgml⁻¹ CaCl₂ and 0.8 mgml⁻¹ thiamine hydrochloride and was then filter sterilised. The two solutions were then combined. All purification steps were the same as for non-labelled EAL preparation (263).

Static IR measurements. IR spectra were measured for EAL and AdoCbl using a Bruker Vertex 80 spectrometer with an XSA external sample compartment. A CaF₂ transmission cell with a 100 μm pathlength was used and signals were acquired in the mid-IR region between 1300 – 1700 cm⁻¹.

TRIR Measurements. TRIR measurements were performed using an ultrafast pump-probe setup which has been previously described (64), with a 10 kHz laser repetition rate

---

6 A full account of the ¹⁵N-labelling of EAL and data illustrating the success of the procedure is provided in Section 8.3
and 100 fs time resolution. EAL was concentrated to approximately 200 μM active site concentration. A pulse energy of 1 μJ was used for excitation. The sample was flowed using a peristaltic pump through a cell containing CaF₂ windows and a 100 μm pathlength, and was placed in a rastering sample holder to avoid photobleaching. Difference spectra were measured at time delays after photoexcitation ranging from 100 fs – 2.5 ns. IR changes were measured in the 1300-1700 cm⁻¹ window by two overlapping 128 pixel detectors with a spectral resolution of approximately 3 cm⁻¹ per pixel. The pixel-wavenumber calibration was achieved by measurement of a pure cis-stilbene standard sample, the IR spectrum of which is characterised.

**Stopped-flow UV-vis.** UV-vis stopped-flow (SF/UV-vis) measurements were made in an Applied Photophysics SX.18 MV-R stopped-flow spectrophotometer. For all experiments, EAL was prepared to approximately 30 μM and the absorbance change accompanying interconversion between cob(III)alamin and cob(II)alamin was monitored at 525 nm. Ethylene glycol was prepared to 800 mM, ethanolamine to 10 mM and 2-aminopropanol to 60 μM. These various concentrations were chosen since this allowed signal changes to occur within the 17 second time window over which SF/FT-IR measurements are acquired.

**SF/FT-IR measurements.** For IR measurements, TgK Scientific stopped-flow sample handling unit syringes were driven by a computer-controlled stepper motor. The reaction solutions were fed through a thermostatted “umbilical” to the mixing chamber, which was part of a custom IR flow cell fitted with CaF₂ windows, and mounted inside a small inert atmosphere chamber (also fitted with CaF₂ windows) in a Bruker Vertex 80 spectrometer. For all SF/FT-IR analyses, a scanner velocity of 320 kHz and a spectral resolution of 1 cm⁻¹ were used, allowing IR spectra from 1300-1800 cm⁻¹ to be acquired every 72 ms over a time course of 17 seconds. A shot volume of 150 μl was used to allow all sample inside the transmission cell to be expelled before the next reading; with full sample syringes this allowed 25 SF/FT-IR readings to be acquired per experiment. All EAL and substrate samples were equilibrated to 25 °C for 10 minutes before analysis, and EAL was prepared to ~ 150 μM for all experiments. Ethylene glycol was prepared to a concentration of 800 mM, ethanolamine to 10 mM and 2-aminopropanol to 400 μM to allow measurement within the acquisition window of the SF/FT-IR. For 2-aminopropanol a larger substrate concentration was prepared relative to stopped-flow UV-vis measurements to account for the higher EAL concentration used for SF/FT-IR. Data manipulation involved discarding the initial 9 shots, as these are within the dead volume of the instrument, then averaging the subsequent readings (~ 15 repeats). Difference spectra were then calculated relative to the first clean spectrum: 144 ms (non-labelled EAL vs. ethylene glycol and ethanolamine), 482 ms (non-labelled EAL vs. 2-aminopropanol) and 277 ms (¹⁵N-labelled EAL vs. ethanolamine).
4.3 Results and Discussion

Our previous TRIR data from free AdoCbl showed a sub-picosecond appearance of signal, which then decays over about 2 ns (Chapter 3, (260)). These data fitted most accurately to the sum of four exponentials, in close agreement with similar UV-vis ultrafast transient absorption (TA) measurements (117). Selected TRIR difference spectra acquired between 1 ps – 2 ns from the EAL holoenzyme (Figure 4.2b) are compared to data from free AdoCbl in Figure 4.2c. Of particular note are the appearance of a ground state bleach at about 1650 cm\(^{-1}\) and a transient signal at about 1661 cm\(^{-1}\) in the EAL measurements, both of which coincide with the protein amide I band and decay with similar kinetics to the signal from the cofactor. There are also differences around 1425–1475 cm\(^{-1}\), which coincide with the amide II bands of deuterated proteins (264). The lower concentration of the protein sample (about 200 µM) compared to the free cofactor (about 8 mM) meant that we had to average more spectra to acquire good quality data, and therefore collected fewer data points. Consequently, the principal kinetics from single value decomposition (SVD) analysis only fit to the sum of two exponentials as opposed to the four exponential fit to the AdoCbl data (see Figure S4.1 and Table S4.1 in the Supplementary Information). Although protein binding is known to affect the kinetics of AdoCbl photolysis (156, 157, 163), the most marked effect is on the excited state processes, which have not been fully resolved here. However, values for \(k_3\) and \(k_4\)—which represent radical pair formation and dynamics—are of the same order across the available TRIR and UV-visible data sets (Table S4.1).

Our SF/FT-IR spectrometer measures spectra every 72 ms, and therefore acquisition of a full data set of the pre-steady states for ethanolamine or 2-aminopropanol (slow substrate) was not possible. Therefore, initial measurements focused on ethylene glycol (EG), a quasi-substrate that initiates homolysis and H abstraction to give cob(II)alamin and the substrate radical (EG\(^{•}\)) over a number of seconds, and produces acetaldehyde as product (P). However, it renders EAL inactive after a single turnover by coordination of a different ligand (X) to the 6th position of the Co centre (Equation 4.1) (265).

\[
\left(\text{Co}^{\text{II}} - \text{Ado}\right) + \text{EG} \xrightarrow{\text{slow}} \left(\text{Co}^{\text{II}}.\text{AdoH}\right) \cdot \text{EG}^{•} \xrightarrow{\text{[} \rightarrow \left(\text{Co}^{\text{III}} - X.\text{AdoH}\right) + P \right]}
\text{ (Eqn. 4.1)}
\]
Figure 4.2. Comparison of AdoCbl and EAL static FT-IR spectra and TRIR difference spectra. (a) Ground-state FT-IR spectra of EAL holoenzyme (solid line) and AdoCbl (red line), normalized to peak absorbance. (b) Selected TRIR difference spectra from EAL holoenzyme. (c) Overlaid 10 ps TRIR difference spectra from the EAL holoenzyme (solid line) and AdoCbl (red line). The EAL spectrum varies in the region around 1450 cm\(^{-1}\), a bleach at 1650 cm\(^{-1}\), and a transient at 1661 cm\(^{-1}\) (A=absorbance).

UV-vis absorbance changes were initially acquired by stopped-flow (SF/UV-vis) at 525 nm and show the conversion of cob(III)alamin to cob(II)alamin over 15 s (Figure S4.2). Reversion to cob(III)alamin (the part of Equation 4.1 in squared brackets) is not observed in this time. Normalised difference spectra from similar SF/FT-IR experiments were calculated by subtracting the first clean spectrum—which is dominated by the intact holoenzyme and free substrate—from the subsequent time points (Figure 4.3a). The most striking spectral features are at 1650 and 1661 cm\(^{-1}\), where the signals evolve at similar rates. An example trace at 1661 cm\(^{-1}\) illustrated in Figure 4.3b was fit to the sum of two exponentials, with rate constants \(k_1 = 1.15 \pm 0.18 \text{ s}^{-1}\) and \(k_2 = 0.27 \pm 0.04 \text{ s}^{-1}\). Interestingly, these peaks correspond closely in position to the additional signals observed in the TRIR data from photolysis of the EAL holoenzyme (Figure S4.3). EAL and AdoCbl both have amide bands in this region, and therefore both the SF/FT-IR and TRIR signals could either represent structural changes within the cofactor or the protein, or interactions between the two. Other notable features include bleaches at 1625 and 1555 cm\(^{-1}\), and a transient at 1540 cm\(^{-1}\).
Figure 4.3. SF/FT-IR analysis of EAL vs. ethylene glycol. (a) Selected SF/FT-IR difference spectra of 150 mM EAL versus 800 mM ethylene glycol. The major signal has peaks at 1661 and 1650 cm$^{-1}$. (b) Example SF/FT-IR trace at 1661 cm$^{-1}$ for 150 mM EAL versus 800 mM ethylene glycol, fitted using the sum of two exponentials.

Although full measurement of pre-steady state data with ethanolamine and 2-aminopropanol (the part of Equation 4.2 set in square brackets) was not possible, there remained the possibility of capturing signal changes at the end of turnover.

\[
\left[ (\text{Co}^{\text{III}} - \text{Ado}) + S \xrightarrow{\text{fast}} (\text{Co}^{\text{II}}.\text{AdoH}) \right] S^* \xrightarrow{\text{slow}} (\text{Co}^{\text{II}}.\text{Ado}) + P \quad (\text{Eqn.4.2})
\]

Absorbance changes at 525 nm after rapid mixing of EAL with excess substrate (S) show a period of steady-state turnover followed by a relatively sharp absorbance change for ethanolamine (about 300 ms), with a more gradual (about 10 s) change for 2-aminopropanol (Figures S4.4 and S4.5, respectively). These correspond to chemical changes between cob(II)alamin/substrate radical pair that accumulates during turnover (256), and the recombination of the Co-C bond upon substrate exhaustion (Equation 4.2) (154, 159). SF/FT-IR difference spectra were again calculated by subtracting the first clean spectrum (144 ms for ethanolamine and 482 ms for 2-aminopropanol) from spectra at subsequent points in time. In contrast to ethylene glycol, these early spectra are dominated by the cob(II)alamin/substrate radical pair that accumulates during turnover. The most intense signals that evolve after the exhaustion of substrate are found at the same frequencies as
those from the EG experiment, but are inverted (Figure 4.4a). Much like the SF/UV-vis data in Figure S4.4, the SF/FT-IR difference spectra with ethanolamine as substrate (Figure S4.6) show a period of steady-state turnover followed by the evolution signals over 300 ms (Figure 4.4b and Figure S4.7). These difference spectra also included two additional signal increases at 1715 and 1354 cm$^{-1}$ that had linear evolution profiles (Figure 4.4a and Figure S4.6), and are due to the formation of acetaldehyde (Figure S4.8) (266). SF/FT-IR spectroscopy of EAL versus 2-aminopropanol yielded similar difference spectra to ethanolamine (Figure 4.4a and Figure S4.9). They differ in that the signal changes appear over about 10 s (Figure 4.4a and Figure S4.10) similar to that observed in the UV-vis region, and the propanal (product) peaks are not visible because of the lower substrate concentration (400 µM, compared to 10 mM ethanolamine). The signal inversion compared to EG is perhaps unsurprising, as the difference spectra in each case are calculated with reference to different early time spectra—the intact holoenzyme for EG, and cob(II)alamin/substrate radical pair for ethanolamine and 2-aminopropanol. Therefore, inversion at 1625, 1650, and 1661 cm$^{-1}$ represent, at least in part, the equivalent but opposing processes, that is, Co-C bond homolysis and recombination. The signals at 1555 to 1540 cm$^{-1}$ are similarly inverted.

To assign the major SF/FT-IR signal changes at 1650 and 1661 cm$^{-1}$ to either the protein or cofactor, we isotopically labelled the EAL apoenzyme with $^{15}$N, which results in a detectable red shift of 2 cm$^{-1}$ of the amide I band (267). EAL was then incubated with non-labelled AdoCbl shortly before data acquisition. Since the EAL-ethanolamine SF/FT-IR experiment yields a large acetaldehyde signal at 1715 cm$^{-1}$, this was used as an internal control that should not shift between isotope experiments. SF/FT-IR data for non-labelled and $^{15}$N-labelled EAL yielded comparable difference spectra (Figure 4.5a). As expected, the acetaldehyde product peak has not shifted (Figure 4.5b). However, the 1650 and 1661 cm$^{-1}$ signals from $^{15}$N-labelled EAL are red-shifted by 2 cm$^{-1}$ (Figure 4.5c and Figure S4.11), which suggests that they are at least in part reporting on changes in the protein. As the amide II band can vary significantly depending on the extent of protein deuteration, the signal at 1555–1540 cm$^{-1}$ was not similarly analysed.
Figure 4.4. Summary of SF/FT-IR analysis of EAL vs. ethanolamine, and EAL vs. 2-aminopropanol.
(a) Overlaid, final time point SF/FT-IR difference spectra relative to the 144 ms spectrum for ethylene glycol (black) and ethanolamine (red), and relative to the 482 ms spectrum for 2-aminopropanol (blue).
(b, c) Example SF/FT-IR traces at 1661 cm$^{-1}$ for 150 mm EAL versus (b) 10 mM ethanolamine and (c) 400 mM 2-aminopropanol. These traces are comparable with the equivalent SF/UV-vis (Figures S4.4 and S4.5).
Figure 4.5. SF/FT-IR spectra for $^{15}$N-labelled and non-labelled EAL vs. ethanolamine. Shown is the superposition of normalised SF/FT-IR difference spectra of non-labelled (black line) and $^{15}$N-labelled EAL (red line). (a) Measured after turnover of 10 mM ethanolamine. (b), (c) Spectra 1.6 s post-mixing showing: (b) no shift in the acetaldehyde product peak at 1715 cm$^{-1}$ and (c) a 2 cm$^{-1}$ red-shift in the peaks at 1650 and 1661 cm$^{-1}$ from $^{15}$N-labelled EAL. The signal variation in (a) between 1400–1600 cm$^{-1}$ may be owing to noise, but can also be attributed in part to peak shifts that result from potentially different extents of deuteration.
4.4 Summary and Conclusions

In conclusion, both TRIR and SF/FT-IR data from studies with EAL show a signal bleach at 1650 cm\(^{-1}\) and transient at 1661 cm\(^{-1}\) (amide I region) that correlates with the reaction chemistry. It is possible that the AdoCbl spectrum shifts on binding to EAL, and that both sets of signal changes in this region result from the coenzyme. However, the red-shift in the SF/FT-IR data after isotopic labelling of the protein strongly suggests that much of the signal results from a structural change in the protein scaffold. The about 3 cm\(^{-1}\) resolution of TRIR prevented a similar \(^{15}\)N-labeling experiment (64). However, the coincidence of these TRIR signals with those from SF/FT-IR spectroscopy (Figure S4.3) is certainly compelling, and may indicate a measurable impact on the EAL protein from Co-C bond photolysis. Such an interaction is consistent with the viscosity dependence of the recombination kinetics after photolysis in EAL, recently attributed to a mobile active-site residue (likely Glu\(^{287}\)) (101) interacting with the adenosyl radical (163). Whether the TRIR peaks at 1650 and 1661 cm\(^{-1}\) are from the protein or otherwise, they are a relatively minor component of the signal, the majority of which resembles the predicted changes in the cofactor that accompany homolysis (calculated using DFT) (260). The fact that such cofactor signals are not clearly evident above the noise in the SF/FT-IR spectra suggest that these data are dominated by signals representing changes in the protein following substrate binding or turnover. The signal bleach at 1650 cm\(^{-1}\) represents a change of approximately 1 % of the overall IR absorbance for the holoenzyme at this wavenumber, which corresponds to about 7 ± 2 peptide bonds in the EAL heterodimer that binds each cofactor (see the Supplementary Information). These changes accompany the reaction chemistry, and appear to be more substantial than those observed in the EAL crystal structure from substrate binding alone (101), where the chemistry was arrested by using a AdoCbl derivative that does not undergo homolysis. The implications of directly monitoring the protein structure by IR methods during the EAL-catalysed reaction are wide ranging, and suggest that such techniques provide a powerful resource for answering long-standing questions regarding AdoCbl-dependent enzymes, and potentially other B\(_{12}\)-dependent processes.
4.5 Supplementary Information

Table S4.1. A comparison of the kinetic parameters from SVD analyses of both TRIR and TA data following photolysis of AdoCbl and EAL-AdoCbl. Owing to the lower concentration of protein sample (~200 μM), alongside the shorter pathlength used in the TRIR studies (100 μm), more averaging was necessary to improve data quality and therefore fewer data points were acquired. As a result, the TRIR data from EAL-AdoCbl were only fit to the sum of 2 exponentials. However, values for $k_3$ and $k_4$ are of the same order across the data sets.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>AdoCbl (TRIR) (260)</th>
<th>AdoCbl (TA) (119)</th>
<th>EAL-AdoCbl (TRIR)</th>
<th>EAL-AdoCbl (TA) (163)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$ ($k_{ex1}$)</td>
<td>1.73 ± 0.225 ps$^{-1}$</td>
<td>0.69 ± 0.05 ps$^{-1}$</td>
<td>-</td>
<td>5.6 ± 0.5 ps$^{-1}$</td>
</tr>
<tr>
<td>$k_2$ ($k_{ex2}$)</td>
<td>0.229 ± 0.012 ps$^{-1}$</td>
<td>0.071 ± 0.001 ps$^{-1}$</td>
<td>-</td>
<td>0.66 ± 0.04 ps$^{-1}$</td>
</tr>
<tr>
<td>$k_3$</td>
<td>17.5 ± 1.54 ns$^{-1}$</td>
<td>9.0 ± 1.0 ns$^{-1}$</td>
<td>15.5 ± 3.44 ns$^{-1}$</td>
<td>15.8 ± 0.8 ns$^{-1}$</td>
</tr>
<tr>
<td>$k_4$</td>
<td>2.32 ± 0.19 ns$^{-1}$</td>
<td>2.0 ± 0.2 ns$^{-1}$</td>
<td>1.99 ± 0.25 ns$^{-1}$</td>
<td>1.21 ± 0.03 ns$^{-1}$</td>
</tr>
</tbody>
</table>

Figure S4.1. EAL-AdoCbl TRIR SVD principal kinetic trace. SVD analysis of the EAL holoenzyme TRIR difference spectra using Surface Xplorer software yielded a single principal spectrum. The associated principal kinetic shown were fitted using the sum of 2 exponentials using Origin 8.0, and the kinetics values compared with the available TRIR and TA data in Table S4.1.
Figure S4.2. SF/UV-vis measurement of EAL vs. ethylene glycol at 525 nm. It has previously been shown that ethylene glycol results in cleavage of the Co-C bond, with the adenosyl radical then abstracting hydrogen from the substrate. However, the enzyme is then rendered inactive by coordination of a different upper axial ligand to the Co, with ultimate conversion to hydroxocobalamin (265). Therefore the decay illustrated represents a single turnover of EAL, with perhaps the beginning of a signal increase representing reversion to cob(III)alamin on coordination of the different ligand.

Figure S4.3. Comparison of EAL TRIR and SF/FT-IR difference spectra. Overlaid difference spectra in the 1600-1700 cm⁻¹ region from EAL vs. ethylene glycol SF/FT-IR (7 s spectrum, solid line) and TRIR (2 ps spectrum, red line). The additional bleach (1650 cm⁻¹) and transient (1661 cm⁻¹) in the EAL holoenzyme TRIR data (compared to free AdoCbl) coincide with the major signals from the SF/FT-IR data.
Figure S4.4. SF/UV-vis measurement of EAL vs. ethanolamine at 525 nm. The first 7 seconds of reaction have been omitted as this covers the steady-state turnover where there is little signal change. The sudden increase in absorbance after 9 seconds represents the recombination of the Co-C bond at the conclusion of steady-state turnover when the substrate is exhausted.

Figure S4.5. SF/UV-vis measurement of EAL vs. 2-aminopropanol at 525 nm. The gradual increase after 4 seconds represents the conclusion of steady-state turnover and recombination of the AdoCbl radical pair. The differences between the ethanolamine and 2-aminopropanol UV-vis data could be explained by the binding affinities of each substrate.
Figure S4.6. SF/FT-IR difference spectra of EAL vs. ethanolamine between 1 and 2 seconds. The major signal changes at 1625, 1650 and 1661 cm\(^{-1}\) are inverted relative to ethylene glycol data, Peaks at 1715 and 1354 cm\(^{-1}\) are due to evolution of acetaldehyde (Figure S4.8).

Figure S4.7. Single wavenumber traces from SF/FT-IR difference spectra of EAL vs. ethanolamine at the frequencies of major signal change. These data show similar signal evolutions as the SF/UV-vis measurements (Figure S4.4) with little change in signal initially, representing turnover, followed by signal change over ~ 300 ms corresponding to radical pair recombination.
Figure S4.8. IR signal changes observed when acetaldehyde is left at room temperature for 10 minutes. The difference spectra from the SF/FT-IR EAL vs. ethanolamine experiment (Figure S4.6) show the presence of large acetaldehyde peaks that increase in amplitude linearly at 1355 and 1715 cm\(^{-1}\) until the conclusion of steady-state turnover. However, these peaks also show a gradual linear decrease in absorbance in the subsequent time points. The above control experiment shows static spectra in the main panel where the black trace represents the first IR reading, and suggests that the decrease in absorbance is due to the 20.2 °C boiling point of acetaldehyde. The inset shows difference spectra relative to the initial reading, and better represents the signal decreases at 1715 and 1354 cm\(^{-1}\).

Figure S4.9. SF/FT-IR difference spectra of EAL vs. 2-aminopropanol between 5 and 10 seconds. The major signal changes at 1625, 1650 and 1661 cm\(^{-1}\) are inverted with respect to the ethylene glycol data, and in the same phase as the ethanolamine data.
Figure S4.10. Single wavenumber traces from SF/FT-IR difference spectra of EAL vs. 2-aminopropanol at frequencies of major signal change. These data show a comparable signal change as the SF/UV-vis measurements (Figure S4.5) with no change in signal initially, followed by gradual signal change over ~10 s corresponding to radical pair recombination.

Figure S4.11. Subsequent time point SF/FT-IR difference spectra for both unlabelled and $^{15}$N-labelled EAL vs. ethanolamine. a) 1.71 second data normalised to 1715 cm$^{-1}$ acetaldehyde product peak illustrating no peak shift. b) 1.71 second data normalised to peak positions 1650 and 1661 cm$^{-1}$ illustrating the 2 cm$^{-1}$ red shift for the $^{15}$N-labelled EAL (red trace) data from the unlabelled EAL (black trace). c) 1.77 second data normalised to 1715 cm$^{-1}$ acetaldehyde product peak illustrating no peak shift. d) 1.77 second data normalised to peak positions 1650 and 1661 cm$^{-1}$ illustrating the 2 cm$^{-1}$ red shift for the $^{15}$N-labelled EAL (red trace) data from the unlabelled EAL (black trace).
Relating the signal changes to number of peptides

In this article we have compared a signal bleach at 1650 cm$^{-1}$ and transient at 1661 cm$^{-1}$ in both the TRIR and SF/FT-IR difference spectra (Figure S4.3) to the signal from the cofactor in each case. Our isotope labelling experiments suggest this bleach and transient relate, at least in part, to changes in the protein, and their amplitude relative to the cofactor signal gives us some indication of the magnitude of the protein change in each case. We thus conclude that the protein change resulting from photolysis of the Co-C bond is subtle, and potentially reflects a degree of vibrational coupling between the protein and AdoCbl chromophore. On the other hand, the protein signal overwhelms the SF/FT-IR data, suggesting a more substantial protein change that accompanies the reaction chemistry after substrate binding. In each case, it would be interesting to approximate the number of peptides ($n$) that relate to the corresponding signal change. A simple method is presented in Equation 4.3 below, and involves calculating the fractional signal change represented by the bleach at 1650 cm$^{-1}$ (change in absorbance divided by total absorbance, $\Delta A_{1650}/A_{1650}$), and then multiplying that by the number of amino acids (751) in the heterodimer that binds each AdoCbl:

$$n = 751 \left( \frac{\Delta A_{1650}}{A_{1650}} \right)$$

(Eqn. 4.3)

Achieving a truly accurate figure is not straightforward, however, especially since both the protein and cofactor contribute to the total IR absorbance in the amide I region. The AdoCbl absorption peak in this region shifts upon binding to EAL, presumably because the corrin amides are involved in binding. The result, however, is that deconvoluting the absorbance contributions from protein and cofactor isn't really possible, because the shift in the AdoCbl spectrum on binding precludes simply subtracting the free cofactor spectrum from that of the holoenzyme. Further, there may not be 100 % cofactor occupancy in the sample. If this is the case, all of the protein present will contribute to the total absorption spectrum, but not to the difference spectrum. However, one can make a crude approximation based on the assumptions of 100 % occupancy and that the sheer number of peptide bonds in the protein compared to the number of amides in the cofactor are likely to mean the protein signal is dominant in this region.

In the case of the SF/FT-IR data with ethylene glycol as substrate, the maximum amplitude for the 1650 cm$^{-1}$ bleach in the difference spectra is $\Delta A_{1650} = 9.06 \times 10^{-3}$, and the total holoenzyme absorbance is $A_{1650} = 0.77$. Using equation 1, this gives a percentage change of ~ 1.2 %, which corresponds to $n$ ~ 9 amino acids. Similar calculations with ethanolamine
(ΔA_{1650} = 5.79 \times 10^{-3}, A_{1650} = 0.922) give 0.6 % and ~ 5 amino acids; and 2-aminopropanol
(ΔA_{1650} = 6.10 \times 10^{-3}, A_{1650} = 0.757) give 0.8 % and 6 amino acids. These don’t seem unrealistic
figures, and the variation is expected from both the approximate nature of the calculations
and the fact the substrates are different (with very different rates, sizes, mechanisms etc.).
They may, therefore, give us a credible idea of the approximate number of residues involved
in catalysis during turnover.

An equivalent calculation for the TRIR data requires further considerations in
addition to those discussed above. If the signals at 1650 cm\(^{-1}\) and 1661 cm\(^{-1}\) do originate from
the protein as a result of vibrational coupling with the chromophore, the quantum yield of
excited states and photoproducts is important. Although in principle it is possible to account
for this, we know nothing of how efficient any vibrational coupling may be. Taken together
with the difficulties relating to separation of the protein and cofactor signals, the figures
reached are less credible than for the SF/FT-IR data. The TRIR bleach amplitude at 1650 cm\(^{-1}\)
is ΔA_{1650} = 3.16 \times 10^{-5}, while the absorbance at 1650 cm\(^{-1}\) taken from a FT-IR spectrum of a
sample of the same concentration is A_{1650} = 1.52. From Equation 4.1 we get a percentage
change of 0.002 %, which corresponds to n = 0.016 amino acids. Although a less satisfactory
result, it does help support our central conclusion that, during the enzyme catalysed reaction,
the protein is more directly and substantially involved in the chemistry, whereas photolysis
alone demonstrates a more modest vibrational coupling between protein and cofactor (163).
5. Modulation of ligand-haem reactivity by binding pocket residues demonstrated in cytochrome c’ over the femtosecond-second temporal range

5.1 Abstract

The ability of haemoproteins to discriminate between diatomic molecules, and the subsequent affinity for their chosen ligand, is fundamental to the existence of life. These processes are often controlled by precise structural arrangements in proteins, with haem pocket residues driving reactivity and specificity. One such protein is cytochrome c’ (Cytcp), which has the ability to bind nitric oxide (NO) and carbon monoxide (CO) on opposite faces of the haem; a property shared with soluble guanylate cyclase. Like soluble guanylate cyclase, Cytcp also excludes O$_2$ completely from the binding pocket. Previous studies have shown that the NO binding mechanism is regulated by a proximal arginine (R124) and a distal leucine (L16) residue. Here, we have investigated the roles of these residues in maintaining the affinity for NO in the haem binding environment by using various time resolved spectroscopy techniques that span the entire femtosecond-second temporal range in the UV-vis, and femtosecond-nanosecond by IR. Our findings indicate that the tightly regulated NO rebinding events following excitation in wild-type Cytcp (WT) are affected in the R124A variant. With R124A, vibrational and electronic changes extend continuously across all timescales (from fs-s), in contrast to WT, and an L16A variant. Based on these findings, we propose a NO (re-)binding mechanism for the R124A variant of Cytcp that is distinct from that in WT. In the wider context, these findings emphasise the importance of haem pocket architecture in maintaining the reactivity of haemoproteins towards their chosen ligand and demonstrate the power of spectroscopic probes spanning a wide temporal range.

5.2 Introduction

Due to their fundamental roles in biochemical pathways as widespread as respiration, vasodilation and drug metabolism (175), haemoproteins continue to attract major interest in relation to mechanisms of diatomic gas binding. Discrimination between NO, CO and O$_2$ by haemoproteins is a remarkable example of biological specificity and one of the most challenging questions at present is understanding how the protein structure regulates the affinity of the haem cofactor towards these diatomic gases (268). Haemoproteins are known to select between these similar diatomic gases in order to modulate their functionality, for

---

7This article was published as:
example as sensory proteins or gas transporters. The mono-His Cytcps are present in a variety of nitrogen-fixing, denitrifying and photosynthetic bacteria. Physiologically, Cytcp has been suggested to be implicated in NO transport and reduction of intracellular NO toxicity (188-191). In a remarkable example of ligand discrimination and specificity, Cytcp excludes O$_2$ completely from the haem binding site; while it binds CO on the distal face forming a 6-coordinate haem, and NO on the proximal face as a 5-coordinate haem (269). The proximal binding of NO as a 5-coordinate adduct is highly unusual, however, a protein that has been hypothesised to share this feature is soluble guanylate cyclase (sGC); a NO sensor that facilitates vasodilation and neurotransmission (175, 178). Previous studies have suggested that sGC binds NO in a two step biomolecular mechanism that yields a proximally-bound 5c-NO complex (270). As no structural data are available for the haem domain of sGC, mechanistic understanding of sGC activation by NO has benefited greatly from studies of Cytcp (270, 271). The crystal structures of Alcaligenes xylosoxidans cytochrome c’ (AxCytcp) in the ferrous and NO-bound forms are shown in Figure 5.1 (211, 217, 220), highlighting the 5c proximally bound histidine (Figure 5.1a) and haem-bound NO (Figure 5.1b). A proposed NO binding mechanism is outlined in Figure 5.2, which is based on a number of spectroscopic and crystallographic studies (211-214, 220). Initially, NO binds distally to form a 6c-NO intermediate, which weakens the proximal His-Fe bond due to the repulsive trans effect of NO binding. This allows displacement of the His residue from the haem by a second NO molecule, forming a putative dinitrosyl intermediate, from which the distal NO dissociates to leave the proximal 5c-NO adduct.

Studies involving an L16A variant of AxCytcp have suggested that the hydrophobic distal pocket (including L16) is crucial to the NO binding mechanism and ligand discrimination (217, 218). With the L16A variant, NO binds distally as a 6c-NO adduct, but is unable to disrupt the His-Fe bond, causing this distal complex to be trapped. Furthermore, in this variant CO has been observed to bind with the highest affinity reported for any haemoprotein (217), and O$_2$ is no longer excluded from the distal face (218). Another residue that contributes to the NO binding mechanism is the Arginine 124 in the proximal haem pocket. When this residue was exchanged for an alanine residue (R124A), the His-Fe-NO intermediate was not detected spectroscopically, suggesting that R124 plays a part in extending the lifetime of this intermediate (220). Also, the crystal structure of this variant revealed a mixture of proximally bound 5c-NO (analogous to the wild-type) and distally bound 5c-NO with the haem shifted "up" into the cavity vacated by Arg$^{124}$ (Figure 5.3) (220). These findings suggest that the stability of haem binding to the protein scaffold is somehow facilitated by R124, and emphasises its importance structurally.
Figure 5.1. The haem binding pocket crystal structures of reduced and NO-bound AxCytcp. (A) Illustrates the ferrous AxCytcp at 1.45 Å resolution (PDB 2yli) (217), and (B) the ferrous, NO-bound AxCytcp at 1.2 Å resolution (PDB 2xlm) (220) with key binding pocket residues shown.
Figure 5.2. **The currently proposed distal to proximal NO binding mechanism of AxCytcp.** Adapted from ref. (211). Initially, an NO molecule binds to the hydrophobic distal face forming a 6c-NO species. The Fe-His bond is then cleaved, allowing a second NO to bind on the proximal face forming a putative dinitrosyl intermediate. Finally, the distal NO detaches, leaving a proximally bound, 5c-NO species. Hypothesized, short-lived intermediates are denoted by square brackets. Arg^{124} is thought to prolong the lifetime of the 6c-NO distal species by steric and/or electrostatic effects on His^{120}.

Figure 5.3. **R124A AxCytcp mixture of NO-bound structures.** The haem environment for NO-ligated R124A shows a mixture of 5c-NO species. The major conformation (orange) is a proximally bound 5c-NO species with two orientations of the bound ligand (occupancies of 0.3 and 0.4), while the minor conformation (blue) has a distally bound 5c-NO with the haem face shifted into the cavity vacated by the proximal arginine residue (0.3 occupancy). (PDB 2xl6, adapted from ref. (220))
UV-vis ultrafast transient absorption (TA) and resonance Raman measurements have shown that wild-type AxCytcp (WT) binds 5c-NO in a highly controlled environment (223, 224). This is demonstrated by the high proportion of geminate recombination after excitation (~ 99 %) on an ultrafast timescale (τ = 7 ps) (223). Laser-flash photolysis studies have shown that only ~ 1 % of the population releases NO to the surrounding bulk solvent with recombination of the proximal histidine to the haem Fe (222). This reformation of the His-Fe bond prevents NO rebinding on the proximal face and has been described as a "kinetic trap" mechanism. Any further NO rebinding cannot be directly to the proximal face but must occur via the 6c-distal intermediate which enables Fe-His bond breakage. These findings highlight the role of the proximal pocket structure in minimising NO escape, with the aforementioned proximal arginine (R124) hypothesised to directly protect against escape to the surrounding solvent (224).

As the control of NO binding is thought to be diminished in the R124A and L16A variants, we have investigated the roles of these residues in controlling the haem pocket environment with respect to NO binding in AxCytcp. We have used pump-probe TA experiments and laser-flash photolysis to continuously sample all relevant timescales from fs-s for NO rebinding. Furthermore, these TA studies have been complemented with ultrafast time-resolved infrared (TRIR) measurements from fs-ns. TRIR is a powerful tool for analysing molecular vibrations following photoexcitation, which provides direct information on bond cleavage and reformation, as well as any protein-heme interactions. Thus, the geminate recombination events observed from fs-ns can be probed by TA (monitoring electronic transitions) and TRIR (monitoring vibrational transitions), in order to yield insights into NO-Cytcp events following photoexcitation. For R124A in particular, this has revealed the crucial role of this residue in modulating haem pocket reactivity for NO rebinding and control of NO escape to bulk solvent.

5.3 Materials and Methods

*AxCytcp preparation.* Wild-type, R124A, and L16A AxCytcp were overexpressed and purified as previously described (217, 219). WT and R124A samples were isolated in MES buffer, pH 6.0 in the ferric form and purity was estimated by SDS-polyacrylamide gel electrophoresis and UV-vis spectroscopy, with the concentration estimated using ε400 = 80,000 M⁻¹cm⁻¹ (272). Samples were reduced using an excess of sodium dithionite (~ 10 mM) in an anaerobic glove box (Belle Technology) and passed down a desalting column (Centri Pure P25 – EMP Biotech) equilibrated with anaerobic 50 mM CHES buffer, pH 8.9 (D₂O for IR

---

8Details of the successful generation of the L16A and R124A variants are detailed in Section 8.4. A full account of the overexpression and purification procedure for Cytcp is detailed in Section 8.5.
measurements adjusted to pD 8.9) to remove excess reductant. The concentration of the ferrous sample was estimated using $\varepsilon_{426} = 97,000 \text{ M}^{-1}\text{cm}^{-1}$. (272). The L16A variant was isolated with CO bound to the haem distal face as previously reported (217). This ligand was removed by incubating with an excess of potassium ferricyanide (500 mM) for one hour at room temperature under anaerobic conditions. Excess oxidant was removed by passage down a desalting column equilibrated with anaerobic 50 mM MES buffer, pH 6.0. The sample was then reduced using an excess of sodium dithionite (~10 mM) and passed down a desalting column pre-equilibrated with anaerobic 50 mM CHES buffer, pH 8.9 (D$_2$O for IR measurements adjusted to pD 8.9). The concentration was estimated using $\varepsilon_{420} = 80,000 \text{ M}^{-1}\text{cm}^{-1}$.

**TA measurements.** A Ti:sapphire amplifier (a hybrid Coherent Legend Elite-F-HE) was pumped by a Q-switched Nd:YLF laser (Positive Light Evolution-30) and seeded by a Ti:sapphire laser (Spectra-Physics Mai Tai). The amplifier has an output wavelength of 800 nm, a 1 kHz repetition rate, and a 120 fs pulse duration. This beam is then split, with part of the output used to produce tuneable radiation in the range 250 to 1000 nm via a non-collinear optical parametric amplifier (Light Conversion TOPAS-White). Another fraction of the amplifier output is used to pump Helios and Eos (Ultrafast Systems LLC) broadband pump-probe transient absorption spectrometers, with instrument response functions of around 170 fs and 500 ps, respectively. In all cases, samples were excited at 532 nm with 0.5 – 1 µJ power and a beam diameter of ~ 200 µm. Absorption changes were monitored between 350 – 700 nm at time delays between 100 fs – 3 ns after excitation for Helios experiments and from 0.5 ns – 400 µs for Eos experiments. Reduced samples were added to 2 mm pathlength quartz cuvettes and adjusted to a concentration such that the Soret band had an absorbance reading of ~ 0.7 (~ 44 µM). A Suba seal was then attached to the cuvette entrance inside an anaerobic glovebox and NO gas bubbled into the cuvette until the spectrum resembled the respective NO-bound profile. Samples were stirred to avoid photobleaching.

**Flash photolysis measurements.** Laser-flash photolysis measurements were carried out as described previously (273), which uses a 150 W xenon arc lamp probe that is pulsed for measurements > 1 ms and adjusted to the probing wavelength by an input monochromator. For saturated solutions (2 mM) of NO, samples were prepared by injecting NO gas into Suba sealed 1 cm pathlength quartz cuvettes containing reduced AxCytcp under anaerobic conditions. The NO-bound Soret absorbance reading was then adjusted to ~ 0.7 (~ 9 µM). An excitation wavelength of 532 nm was generated using the second harmonic of a Q-switched Nd:YAG laser (Brilliant B, Quantel) with a laser energy at the sample of 100 mJ and beam diameter ~ 1 cm. Five replicates were acquired per sample and the temperature maintained at 25 °C. The detection system for ms-s measurements utilized a photomultiplier.
tube, whilst faster μs-ms measurements were acquired using an Infinium oscilloscope model no. 54830B.23 For NO concentration-dependence studies of the R124A variant, a saturated solution of NO was prepared by adding anaerobic 50 mM CHES buffer, pH 8.9 to a 50 ml conical flask followed by attachment of a Suba sealed. NO gas was then injected into the conical, replacing 25 ml of buffer and mixed thoroughly. This NO solution was diluted appropriately by injection into a Suba sealed cuvette containing R124A and a varying volume of anaerobic buffer. These samples were analysed shortly afterwards to prevent fluctuation in NO concentration, and 7 replicates were acquired to calculate shot-to-shot error.

**Stopped-flow UV-vis spectroscopy.** For SF/UV-vis measurements, WT and R124A were prepared to a concentration of 12 μM (6 μM post-mixing), while L16A was prepared to 4 μM (2 μM post-mixing) due to its higher affinity for NO. For NO concentration-dependence measurements, a stock NO solution (2 mM) was prepared in a 50 ml conical flask as for laser-flash photolysis. Under anaerobic conditions the NO stock solution was diluted appropriately to the desired concentrations with 50 mM CHES buffer, pH 8.9. Photodiode array (PDA) and single wavelength absorption measurements were acquired using a TgK Scientific stopped-flow spectrometer housed inside an anaerobic glovebox (Belle Technology). Samples were rapidly mixed together to trigger NO binding and 6 readings were acquired to calculate shot-to-shot error.

**TRIR spectroscopy.** TRIR measurements were carried out using an ultrafast TRIR experimental setup described previously (64), with a 10 kHz repetition rate and 100 fs time resolution. AxCytcp samples were prepared to a ferrous form at a concentration of ~ 2 mM, then either analysed without NO bound or incubated with 3.5 mg spermine NONOate (Tocris bioscience – Cat No. 1135) for 4 hrs at room temperature to allow NO binding. Full NO binding to ferrous Cytcp samples was confirmed by UV-vis spectroscopy (Figure S5.1) Samples were then added to an anaerobic cell with CaF₂ windows and a 75 μm spacer, which was rastered to avoid sample damage. For all samples an excitation wavelength of 532 nm was used with 1 μJ pulse power and a beam diameter of ~ 150 μm set at the magic angle with respect to the IR probe beam. Spectra were measured at time delays ranging between 500 fs and 1 ns. Difference spectra were generated relative to the ground state in the spectral window 1300-1800 cm⁻¹. This was measured using two 128 pixel detectors with a spectral resolution of approximately 3 cm⁻¹ per pixel. Pixel to wavenumber calibration was calculated as previously shown (260).

**Global fitting.** Kinetic analyses were carried out using Origin 8.5 software. Five spectral positions of significant absorption change in the TA and TRIR were selected and fit with shared lifetime values using the global fitting option. This returns a single lifetime value for each exponential and a number of amplitude values for each of the respective frequencies.
The sample concentration and pump power was similar in all the TA experiments performed, however, when combining the fs-ns and ns-µs datasets to span the full fs-µs time range, a small scaling factor (1.2) was applied to the ns-µs dataset to ensure a smooth match to the fs-ns dataset.

5.4 Results and Discussion

5.4.1 Wild-type AxCytcp

Initially TA and TRIR analysis of WT was performed in order to benchmark our findings against existing data, and also to provide previously unreported vibrational spectroscopy data. For both TA and TRIR experiments a similar excitation wavelength (532 nm), laser power (1 µJ) and beam diameter (~ 150 µm) was used, therefore the respective spectral and kinetic data can be directly compared. Furthermore, these experimental conditions were also used in previous ultrafast studies of WT (223, 224), allowing additional comparative analysis. The TA difference spectra are shown in Figure 5.4a with selected data between 0.5 ps and 30 ps, which illustrates the same spectral features as previously observed (223, 224). These include a ground state bleach at 393 nm and an equivalent transient feature centred at 420 nm, both of which decay almost completely to the ground state over ~ 100 ps (for ground state UV-vis absorption spectra of WT and variants see Figure S5.1). The minimal signal remaining after 100 ps (~ 5 % of the initial signal remaining at 396 nm) has previously been assigned to the 5c-His species formed upon NO escape to the surrounding solvent (224).

Figure 5.4b illustrates TRIR difference spectra at selected time intervals after excitation equivalent to the TA spectra in Figure 5.4a (for ground-state IR absorption spectra of WT and variants see Figure S5.2). A number of features are evident in this spectral region, the most prominent of which include signal bleaches at 1656, 1678 (shoulder) and 1577 cm⁻¹, and transient features at 1717, 1745 (shoulder), 1637, 1602 and 1556 cm⁻¹. There are also a number of less pronounced features from 1300-1500 cm⁻¹. The 1656 cm⁻¹ bleach has previously been assigned to the 5c-NO species (212) and therefore provides the opportunity to directly monitor NO geminate recombination. The overall TRIR difference spectra suggest that the vast majority of photolysed NO in WT undergoes geminate recombination on a fast timescale, as previously suggested (223).
Figure 5.4. NO-bound WT AxCytcp TA and TRIR difference spectra. The TA (A) and TRIR (B) difference spectra are relative to the ground state for WT following photolysis of the bound NO by laser excitation. TA difference spectra from 0.5 – 30 ps illustrate a ground state bleach at 393 nm and an equivalent transient at 420 nm which return to the ground state almost completely after 30 ps. The most noteworthy feature in the TRIR difference spectra is the 1656 cm\(^{-1}\) signal bleach corresponding to cleavage of the 5c-NO bond.

In order to determine which spectral features correspond to haem vibrations following excitation, control TRIR experiments were performed using reduced WT, R124A and L16A in the absence of NO. As expected, the difference spectra for reduced WT exhibits a markedly reduced 1656 cm\(^{-1}\) band, but retains a number of other spectral features (Figure S5.3). When compared directly (Figure S5.4), the TRIR data for WT in the presence and absence of NO have a number of common spectral features, particularly in the fingerprint region of 1300-1500 cm\(^{-1}\). We have therefore assigned these signals below 1600 cm\(^{-1}\) to excitation of the heme and structural changes in coordinating protein residues, including the covalently-linked Cys\(^{116}\) and Cys\(^{119}\) (274). It is noteworthy that the 1573 cm\(^{-1}\) band has been previously suggested to report on either the histidine or arginine residue during NO binding.
Our data make either improbable considering this signal bleach is present in the WT with NO bound (histidine displaced), in addition to reduced samples of WT, R124A and L16A (histidine attached, and in the case of R124A, arginine absent) (Figure S5.5).

Major TRIR features between 1600-1800 cm\(^{-1}\) (other than the assigned 1656 cm\(^{-1}\) 5c-NO bleach) include transient absorption features at 1634 and 1717 cm\(^{-1}\). The 1717 cm\(^{-1}\) band could correspond to the emergence of carbonyl (C=O) groups following excitation, as these groups give strong absorbance between 1670 – 1820 cm\(^{-1}\) due to a stretching motion (275). According to the WT NO-bound crystal structure (220), His\(^{120}\) is hydrogen bonded to an aspartate residue (Asp\(^{121}\)) in contrast with the reduced structure. Therefore, following excitation of NO, this His-Asp H-bond could potentially cleave, allowing competition between the His and NO for proximal binding, and temporarily reveal the carbonyl group of Asp\(^{121}\). If this hypothesis were proved correct, this suggests that upon NO-heme photolysis, the Asp-His hydrogen bond is immediately cleaved. However, this band is particularly broad for a single carbonyl stretch, and the possibility of other contributing vibrational features has not been ruled out. The cleavage of the Asp-His H bond also offers an explanation for the transient feature at 1634 cm\(^{-1}\), as histidine has been previously shown to absorb IR about this frequency (276). A summary of our TRIR assignments for WT following excitation is provided in Table S5.1.

Kinetics for the TA and TRIR spectra were calculated by globally fitting the signal decays at five distinct wavelength or wavenumber values, using shared lifetimes and a non-linear least squared fitting model (WT global fitting shown in Figures S5.6 and S5.7). This returned three spectral components as opposed to the previously reported two (Table 5.1) (224), the ~ 7 ps component, which corresponds to geminate recombination of NO, and the ~ 100 ps component corresponding to His rebinding are also observed in our data. The variation in \(\tau_3\) between the TA experiments is likely due to the low amplitude of this lifetime component (with the consequent large error). The TRIR experiments were used principally to determine vibrational changes during geminate recombination (initial 10 ps after excitation), therefore very few data points at longer times were acquired, which explains the poorly resolved \(\tau_3\) value for these data. Since the 1656 cm\(^{-1}\) band of the TRIR has been previously assigned to the Fe-NO bond (212); this allows the geminate rebinding of NO to be directly monitored. This is particularly beneficial in determining the identity of the previously unresolved fast component (\(\tau_1\)). For the majority of NO-binding haemoproteins geminate recombination occurs monoexponentially over ~ 7 ps (277), with any faster components owing to vibrational relaxation of haem after laser excitation (278-280). However, the NO signal bleach in the TRIR (1656 cm\(^{-1}\)) suggests that this fast component also corresponds to NO rebinding, since on these timescales the signal amplitude is returning to the ground state,
as shown in Figure S5.8. The first solved crystal structure of WT suggested that multiple conformers of the NO–haem complex exist, and therefore this component could be fast recombination of a proportion of NO in a different conformation (211). However, this hypothesis must be treated with caution since a later crystal structure reported only one orientation of bound NO (220). It is noteworthy that this fast component is unlikely to correspond to a vibrationally excited NO rebinding to the haem, as no obvious bandshifts are observed, particularly of the 1656 cm⁻¹ feature (281, 282). The 1717 cm⁻¹ feature, which we earlier assigned to Asp¹²¹ in the haem pocket, fitted to a single exponential with τ₁ = 8.15 ± 0.21 ps, which approximately corresponds to the rate of NO geminate recombination. Therefore, it remains possible that this Asp¹²¹ residue gives rise to IR absorbance, until NO undergoes geminate recombination and the Asp-His H-bond is reformed.

**Table 5.1. Global fitting kinetic lifetimes of WT AxCytcp from TA and TRIR data.** These kinetic lifetimes of WT from TA and TRIR data (Figure 5.4a and 5.4b respectively) have been calculated by global fitting to the sum of three exponentials. These lifetime values were generated by global fitting to signal decays at the following frequencies: 377, 393, 417, 428 and 440 nm for the TA data; 1577, 1596, 1625, 1637 and 1655 cm⁻¹ for TRIR data. The τ₃ value in the TRIR was poorly resolved due to the low quantity of data points acquired beyond 20 ps. These values are compared to existing kinetic data from ref. (224).

<table>
<thead>
<tr>
<th></th>
<th>TA</th>
<th>TRIR</th>
<th>Literature TA (224)</th>
</tr>
</thead>
<tbody>
<tr>
<td>τ₁</td>
<td>2.19 ± 0.18 ps</td>
<td>1.31 ± 0.16 ps</td>
<td>-</td>
</tr>
<tr>
<td>τ₂</td>
<td>5.65 ± 0.15 ps</td>
<td>6.13 ± 0.29 ps</td>
<td>7 ± 0.5 ps</td>
</tr>
<tr>
<td>τ₃</td>
<td>204 ± 56 ps</td>
<td>&gt; 20 ps</td>
<td>100 ± 10 ps</td>
</tr>
</tbody>
</table>

Additional laser photoexcitation experiments over longer timescales revealed no further spectral changes on the ns-µs timescale for WT. Hence, the geminate recombination of NO to WT is complete on the fs-ns timescale, suggesting a highly controlled NO rebinding mechanism following photolysis, consistent with the crowded nature of the proximal pocket. However, laser flash photolysis data (µs-s) show that the remaining 5% of signal amplitude in the TA measurements represents NO escape from the haem binding pocket, which subsequently rebinds from the bulk solvent on a much slower timescale (Figure S5.9). The increase in absorbance at 396 nm, which represents formation of 5c-NO species, could be fit to the sum of two exponentials with lifetimes (2.7 ± 0.1) ms and (76.7 ± 0.2) ms. If one assumes saturating NO concentrations in solution (2 mM), these lifetime values roughly equate to reported k₉₉ rates for the distal-to-proximal binding mechanism from stopped-flow experiments (literature values 4.4 × 10⁴ M⁻¹s⁻¹ and 8.1 × 10³ M⁻¹s⁻¹; flash-photolysis values 1.85 × 10⁵ M⁻¹s⁻¹ and 6.5 × 10³ M⁻¹s⁻¹) (213).
5.4.2 L16A AxCytcp

Previous studies have shown that L16A binds NO on the distal face in a 6c-NO mode (with the proximal histidine still attached) (217), without proceeding to a proximal binding mode. Consequently, this variant of Cytcp was a useful comparative system. TA and TRIR difference spectra are illustrated in Figure 5.5, both of which are distinct from the WT spectra. The TA difference spectra exhibits a ground state bleach centred at 416 nm, which corresponds to the ground state UV-vis absorption spectrum of 6c-NO (Figure S5.1), and transient features at 370 and 435 nm that decay to the ground state within 1 ns. The TRIR data show obvious spectral differences to WT (Figure S5.10), with the ground state bleach shifted from 1656 to 1629 cm⁻¹. This is in agreement with an earlier assignment of a 6c-NO species at this position (212), and therefore likely represents the release of NO from the 6c distal binding site, presumably leaving a 5c-His intermediate. The broad transient feature at 1717 cm⁻¹, which was present in the WT sample, is missing in the L16A sample. This complements our hypothesis stating that the > 1700 cm⁻¹ features report on the Asp¹²¹ carbonyl group, since this residue is no longer hydrogen bonded to His¹²⁰ when NO is bound in the 6c distal position. It appears from the TRIR difference spectra that the ground state molecule is fully reformed after ~ 1 ns, with no escape of NO to the bulk solvent and is confirmed by the lack of any signal on the µs – s timescales. This is not a surprising finding as any NO released from the distal pocket into solvent could simply rebind in the absence of any kinetic trap mechanism (as proposed with WT). This is reflected in the kinetics generated from global fitting of the TA and TRIR data, which fitted to the sum of two exponentials. As with WT, this could correspond to two conformations of NO binding distally, but in the absence of an L16A crystal structure this cannot be confirmed. In addition, the lifetime values for L16A appeared to be NO concentration dependent, which represents an unusual feature when considering geminate recombination processes. Due to these anomalous findings, the L16A TA and TRIR data have been used for qualitative comparison with WT and R124A, as opposed to quantitatively. The L16A lifetime values and further discussion is provided in the supplementary information (Table S5.2).
Figure 5.5. NO-bound L16A AxCyt cp TA and TRIR difference spectra. Illustrated are the L16A TA (A) and TRIR (B) difference spectra at equivalent time points following excitation between 1 ps and 1 ns. TA difference spectra exhibit a ground state bleach at 416 nm, in accordance with the ground state spectrum and transients at 370 and 435 nm. The TRIR illustrates a signal bleach at 1629 cm$^{-1}$ which likely corresponds to loss of a 6c-NO species, whilst the majority of other spectral features are also present in the WT TRIR spectra.
5.4.3 R124A AxCytcp

The TA and TRIR difference spectra for the R124A variant at selected time points between 1 ps and 1 ns are shown in Figure 5.6. The TA difference spectra (Figure 5.6a) show a more complex electronic signal than for WT, with increases in absorbance at 414, 426 and 440 nm, all of which have different amplitudes and decay back to the ground state at varying rates. As expected, the ground state bleach remains at 393 nm, which corresponds to the loss of 5c-NO. The increase in spectral complexity is possibly a result of R124A existing as a mixture of proximal and distal 5c-NO species prior to excitation, consistent with the crystal structure (with occupancies of 0.7 for the haem with proximal NO and 0.3 for the haem with distal NO) (220). Furthermore, R124A appears to undergo a lower proportion of geminate recombination compared to WT (~ 15 % of the initial signal amplitude remaining after 1 ns). This increase in residual signal after 1 ns could be due to enhanced solvent exposure at the haem proximal face when the large, basic Arg124 is replaced by an Ala residue; alluding to a role for Arg124 in protecting against NO solvent escape. However, this spectral feature could instead be due to the existence of a mixture of proximally and distally bound NO in the case of R124A.

In contrast to the TA findings, the TRIR data for R124A are comparable to those observed for WT (Figure S5.11). The only exceptions are minimal shifts in peak positions from 1577 to 1573, and 1637 to 1632 cm⁻¹. According to our peak assignments for WT (Table S5.1); the 1573 cm⁻¹ shift is likely due to alterations in the haem-protein vibrational changes following excitation, possibly owing to the mixture of proximally and distally bound 5c-NO for R124A. This characteristic could also explain the peak shift at 1632 cm⁻¹; which we assigned to the loss of His-Ala H-bond and could be affected vibrationally by the mixture of NO binding. As the overall R124A TRIR spectral profile resembles that of WT, this suggests that following excitation the 5c-NO geminate recombination events are vibrationally similar, but have different electronic characteristics.

A significant proportion of the bound NO does not undergo rapid geminate recombination on the fs-ns timescale for the R124A variant and additional spectral changes can be observed on the ns- μs timescale (Figure 5.7). Over the subsequent 10 μs there is a significant increase in amplitude at 432 nm, confirming that there is an additional process occurring with this variant. The absorption band at 432 nm has been previously assigned to the 5c-His species (224), and therefore suggests an increase in this population over μs time scales.
Figure 5.6. NO-bound R124A AxCytcp TA and TRIR difference spectra. Illustrated are R124A TA (A) and TRIR (B) at selected time points between 1 ps and 1 ns. The TA difference spectra exhibit clear differences with those for WT, with transient features at 414, 426 and 440 nm and an enhanced spectral profile at 1 ns. In contrast, TRIR difference spectra exhibit broadly similar spectral features to WT.

Figure 5.7. UV-vis absorption difference spectra for NO-bound AxCytcp R124A from 2 ns – 10 µs. The 2 ns spectrum is similar in profile to the end of the ultrafast (fs-ns) experiments, with a signal bleach at 393 nm and corresponding positive transient absorption feature at 425 nm. An increase in absorbance with time is most clearly seen at 432 nm.
In order to examine this TA data in their entirety, the data on the fs-μs timescales were combined, as described in the Materials and Methods section, to allow kinetic analysis of the whole dataset. Global analysis returned five spectral components instead of three observed with WT, as detailed in Table 5.2. For illustrative purposes, the spectral change at 432 nm is shown in Figure 5.8a, which shows the initial fall in amplitude over ~1 ns, followed by a subsequent rise in absorbance over the following ~10 μs. The early lifetime values for R124A (τ1-τ3) are similar to those for WT, showing two fast phases (thought to be geminate recombination of NO; as with WT the 1656 cm⁻¹ feature is relaxing to the ground state on the ~1 ps timescale) and a ~100 ps feature representing His rebinding to the 4c-haem. For R124A the respective amplitudes of the individually fitted frequencies are larger than for WT and therefore global fitting returned a lifetime with reduced error, and increased similarity with the predicted 100 ps value (224). The increase in τ3 signal amplitude for R124A could be due to additional His rebinding, and therefore in the WT Arg124 may act to repel His120 by virtue of charge and steric effects.

As R124A exists as a mixture of proximal and distal 5c-NO, we suggest that the longer components correspond to the reattachment of 5c-His to the (originally) distal 5c-NO species. This differs from the components present in the distal 6c-NO complex produced at low [NO] in WT where photodissociation of NO leaves a 5c-His species (218). In the R124A case, there is a shifted haem with the proximal histidine stacked parallel to the haem plane. In order for the His to reattach, a significant structural readjustment of the haem is necessary following NO dissociation, which is likely to occur over longer time courses than the originally hypothesized kinetic trap mechanism for WT. We therefore infer that the longer components over ~300 ns correspond to this process, which eventually results in formation of a proximally-bound 5c-His species. The crystal structure of ferrous R124A supports this idea, as only 5c-His is observed in the reduced form (220), analogous to the WT reduced structure. Accordingly, all of the different R124A species that have undergone NO solvent escape will form an identical 5c-His structure after ~300 ns. The proposed mechanism for R124A photolysis on these timescales is outlined in Figure 5.9. This hypothesis relies on the existence of a relatively long lived, highly reactive 4c-haem species, and therefore awaits confirmation, for example by molecular dynamic simulation modeling.
Table 5.2. Kinetic lifetime values from global fitting of TA and TRIR data for R124A across fs-µs (TA) and fs-ns (TRIR) timescales. These lifetime values were generated by global fitting to signal decays at the following frequencies: 393, 402, 416, 427, 431 and 440 nm for TA data; 1575, 1597, 1626 and 1655 cm\(^{-1}\) for TRIR data. The \(\tau_3\) value in the TRIR was poorly resolved due to the low quantity of data points acquired beyond 20 ps.

<table>
<thead>
<tr>
<th></th>
<th>R124A TA</th>
<th>R124A TRIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau_1)</td>
<td>2.65 ± 0.24 ps</td>
<td>0.99 ± 0.12 ps</td>
</tr>
<tr>
<td>(\tau_2)</td>
<td>6.94 ± 0.24 ps</td>
<td>5.15 ± 0.17 ps</td>
</tr>
<tr>
<td>(\tau_3)</td>
<td>123 ± 7 ps</td>
<td>&gt; 20 ps</td>
</tr>
<tr>
<td>(\tau_4)</td>
<td>17.5 ± 2.9 ns</td>
<td>-</td>
</tr>
<tr>
<td>(\tau_5)</td>
<td>302 ± 70 ns</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 5.8. UV-vis 432 nm signal decay following photolysis of 5c-NO R124A from 1 ps – 100 ms. (A) Normalised absorbance measured from ps to µs fitted to the sum of five exponentials. The data acquired from the ps-ns and ns-µs experiments were normalized by their respective absorbance amplitudes at 1 ns – 3 ns. (B) Laser flash photolysis signal over 1 µs – 100 ms fit to the sum of two exponentials. Single wavelength data were acquired in order to achieve an optimal signal to noise ratio. The reduction in noise after ~ 1 ms is due to different detection systems employed. The amplitude of \(\Delta A\) between 1 µs and 10 µs (the overlapping region of the TA and flash-photolysis experiments) has been verified in the Supplementary Material.
Figure 5.9. The proposed geminate recombination mechanism for R124A. This scheme begins with a mixture of proximally and distally bound 5c-NO, with the distally bound form having a shifted haem (denoted by grey shading). Upon excitation both species undergo cleavage of Fe-NO forming a 4c-haem intermediate. Over the subsequent ~ 7 ps geminate recombination occurs for a proportion of each species, which according to the TA data, appears to be diminished with respect to the WT data, with both R124A species likely to contribute to greater NO solvent escape. The remaining proximally bound WT-like species undergoes histidine rebinding over ~ 100 ps forming a 5c-His species, while the distally bound species must undergo haem rearrangement before histidine reattachment, which occurs over ~ 300 ns.
A significant absorbance peak remains at 432 nm for the R124A variant after 10 µs, but subsequent laser-flash photolysis experiments have shown that this decreases on the µs-s timescale (Figure 5.8b). Moreover, analogous stopped-flow UV-vis measurements have shown that, in contrast to the wild-type, the 6c-NO distal species is not observed in the R124A variant. Instead, only the 395 nm 5c-NO and 423/432 5c-His species are observed with a clear isosbestic point at 408 nm (Figure S5.12). This has been ascribed to an increased $k_{6-5}$ in the absence of the effects of an arginine at residue 124. A single exponential fit at increasing NO concentrations provides a $k_{on}$ value of $(6.42 \pm 0.09) \times 10^4 \text{M}^{-1}\text{s}^{-1}$ (Figure S5.13; reported literature value = $(3.2 \pm 0.1) \times 10^4 \text{M}^{-1}\text{s}^{-1}$) (220). In contrast to the stopped-flow findings, the signal decay from the laser photoexcitation measurements in Figure 5.8B fits to the sum of two exponentials and reveals a previously unreported fast phase on the ~ 100 µs timescale. To determine whether this was a second order process, laser-flash photolysis measurements were repeated at a range of NO concentrations for the R124A variant and showed that both $k_1$ and $k_2$ are dependent on NO concentration (Figures S5.14 and S5.15). $k_2$ has a calculated $k_{on}$ of $(7.23 \pm 0.92) \times 10^4 \text{M}^{-1}\text{s}^{-1}$ and is therefore likely to represent the same binding event previously reported in the stopped-flow studies (formation of a 5c proximal NO species). However, $k_1$ has a calculated $k_{on}$ value of $(3.50 \pm 0.15) \times 10^6 \text{M}^{-1}\text{s}^{-1}$, which had not been noticed in previous stopped-flow experiments (the $k_{off}$ value was negligible in both cases). However, at low concentrations of NO (0.1 mM), the signal decay of stopped-flow R124A also fitted to the sum of two exponentials with a rate of decay similar, within an order of magnitude, to the laser-flash photolysis experiments. Due to the time resolution of the stopped-flow instrument (~ 1 ms) this fast phase fitted to only a few data points, hence an increase in error between the two. As this rebinding event from solvent is also NO-dependent, it could correspond to the formation of the distal 5c-NO species (including formation of an initial 6c-NO species, dissociation of histidine, and collapse of the haem structure as an alternative pathway to that which would yield proximal NO).
5.5 Concluding Remarks

The ability of cytochrome c' to bind NO and CO on opposite faces of the haem, and the availability of variants where ligand binding is radically altered, renders it an excellent model system for understanding how precise structural arrangements modulate the affinity of the haem cofactor for these diatomic gases in haemoproteins. By using a variety of time resolved spectroscopic techniques, spanning timescales from fs-s, we have shown the importance of the haem pocket architecture in regulating the affinity for NO. Previous stopped-flow, TA and crystallographic studies have indicated that cytochrome c' binds NO in a tightly regulated manner, with a number of implicated haem pocket residues (211, 213, 220, 222-224). Our findings indicate that removal of a single proximal residue (R124) significantly affects the tightly regulated NO rebinding events upon excitation in WT. In an R124A variant of cytochrome c' there is a significantly different mechanism for haem-NO rebinding, with an increase in NO escape to solvent and an extension into the µs window, which is not observed for either the WT or L16A variant. This is likely to represent a readjustment of the haem cofactor for the distal 5c-NO species and proximal rebinding of His120 to re-form the starting 5c-His species. In addition, there are two distinct, slower processes in R124A only, which likely corresponds to haem collapse and formation of a distal 5c-NO species (fast phase) and formation of a proximal 5c-NO species (slow phase). These results indicate the complexity of NO rebinding in the R124A variant and illustrate the crucial role of this residue in controlling the haem pocket reactivity towards NO.

In the wider context, the ability to monitor ligand photolysis in proteins from initial chemical events (fs) to the final equilibrium position (s) provides an exciting opportunity to monitor protein dynamics corresponding to ligand binding processes. Furthermore, the importance of haem pocket architecture in modulating the control over haem-ligand reactivity has been emphasised, which could have significant implications for other gas-binding proteins.
Figure S5.1. UV-vis spectra for reduced and NO-bound samples of WT, R124A and L16A AxCytcp. These spectra are normalised to $A_{700}$. When reduced, WT (top, left) exhibits a split Soret band with peak positions 424 and 432 nm which shifts to a single Soret upon NO-binding centred at 395 nm. R124A (top, right) has the same peak positions as the wild-type, with the only difference being the reduction in difference between peak intensities in the reduced 424/432 split Soret. L16A (bottom, middle) also has a split Soret in the reduced form with peak positions 420 and 432 nm which upon binding NO exhibits a sharp increase at 416 nm, corresponding to the formation of a 6c-NO species.
Figure S5.2. FT-IR spectra for reduced and NO-bound WT, R124A and L16A AxCytcp, with difference spectra inset. The WT and L16A spectra were normalised to frequencies of no signal change while the R124A was left non-normalised. In all cases the spectra illustrate characteristic protein spectral structure, with an amide I band centred at 1645 cm\(^{-1}\) and amide II centred at 1455 cm\(^{-1}\). The amide II band appears to broadly increase in all cases upon addition of NO, while the amide I stretch shows more subtle changes. For WT (top, left) there are a number of increases which complement existing IR assignments for NO binding to cytochrome c’ (212). R124A (top, right) illustrates similar difference spectra to the wild-type, the main difference being a more predominant IR increase at 1655 cm\(^{-1}\). L16A (middle, bottom) returned difference spectra with an increase at 1630 cm\(^{-1}\) and corresponding decrease at 1655 cm\(^{-1}\) reflecting the 6-coordinate binding event for this variant.
**Figure S5.3. Reduced WT TRIR difference spectra between 1-100 ps after excitation.** This illustrates signal bleaches at 1575, 1471 and 1405 cm\(^{-1}\), and transient features at 1377, 1436 and 1546 cm\(^{-1}\), with a small but markedly reduced signal bleach at 1656 cm\(^{-1}\) which corresponds to a lack of 5c-NO in this species.

**Figure S5.4. Comparison of 1 ps TRIR difference spectrum of WT with and without NO bound.** In the 1300-1600 cm\(^{-1}\) region the majority of signals are identical in position and amplitude. The major differences are the size of the 1655 cm\(^{-1}\) bleach and 1676 cm\(^{-1}\) shoulder, in addition to transient features centred at 1717 and 1637 cm\(^{-1}\).
Table S5.1. Assignment of TRIR difference spectra features for WT AxCytcp.

<table>
<thead>
<tr>
<th>TRIR Absorption Band (cm(^{-1}))</th>
<th>Assignment</th>
<th>IR Absorption Frequency in D(_2)O (cm(^{-1}))</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1710-1750</td>
<td>Asp121, ν(C=O)</td>
<td>1716-1775</td>
<td>(283)</td>
</tr>
<tr>
<td>1656</td>
<td>5c-NO-Heme</td>
<td>1656</td>
<td>(212)</td>
</tr>
<tr>
<td>1634</td>
<td>His, ν(C=C)</td>
<td>1623</td>
<td>(276)</td>
</tr>
<tr>
<td>1300-1600</td>
<td>Heme vibrational modes after excitation</td>
<td>1300-1600</td>
<td>(274)</td>
</tr>
</tbody>
</table>

Figure S5.5. Comparison of TRIR reduced samples for WT, R124A and L16A 1 ps spectra. These spectra are normalised to maximum and minimum amplitudes. These spectra exhibit broadly identical spectral features, the most important of which being the presence of the 1575 cm\(^{-1}\) bleach, which suggests this is unlikely to correspond to the proximal arginine residue, or indeed the histidine residue (this is also present in the NO-bound spectra – Figure S5.2).
**Figure S5.6.** Global fit of TA signal decay for NO-bound WT AxCytcp. The WT TA absorption at five distinct wavelengths have been fitted to the sum of three exponentials using shared lifetimes and a non-linear least squared fitting model.

**Figure S5.7.** The TA residual plot of WT AxCytcp NO rebinding kinetics following photoexcitation. These residual plots are of the 393 and 428 nm traces fitted using the global fitting model shown in Figure S5.6.
Figure S5.8. 1656 cm⁻¹ kinetic decay from WT TRIR data fitted to the sum of two exponentials. These kinetics were comparable to those reported by global fitting in Table 5.1, and the signal decay at time points < 1 ps indicate that the calculated τ₁ value by global fitting (2.19 ± 0.18 ps for TA; 1.31 ± 0.16 ps for TRIR) reports on NO geminate rebinding.

Figure S5.9. Laser-flash photolysis data at 396 nm for WT AxCytcp. This trace has been fitted to the sum of two exponentials.
Table S5.2. The influence of NO / protein concentration on L16A AxCytcp NO rebinding kinetics. Despite these rates reporting on geminate recombination processes, it appears they are dependent on the NO:protein concentration ratio, particularly for $\tau_1$. It is likely that $\tau_1$ corresponds to heme vibrational relaxation events; however, for this to be markedly affected by the concentration of NO and L16A is unusual and may require further experimentation for elucidation.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>[L16A]</th>
<th>[NO]</th>
<th>$\tau_1$</th>
<th>$\tau_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TA</td>
<td>50 µM</td>
<td>100% (2 mM)</td>
<td>1.21 ± 0.21 ps</td>
<td>15.54 ± 0.89 ps</td>
</tr>
<tr>
<td>TA</td>
<td>50 µM</td>
<td>Determined by NONOate – low</td>
<td>3.85 ± 0.17 ps</td>
<td>18.45 ± 0.42 ps</td>
</tr>
<tr>
<td>TRIR</td>
<td>2 mM</td>
<td>Determined by NONOate – low</td>
<td>5.85 ± 0.59 ps</td>
<td>24.6 ± 2.2 ps</td>
</tr>
</tbody>
</table>

Figure S5.10. Comparison of TRIR WT and L16A 1 ps difference spectra. Broadly identical spectral features are observed from 1300-1600 cm$^{-1}$, which most likely corresponds to haem vibrations. The major spectral differences are the shift in bleach from 1656 to 1629 cm$^{-1}$ which correlates to 5c-NO and 6c-NO binding respectively, and the loss of transient signal at 1717 cm$^{-1}$. 
Figure S5.11. Comparison of TRIR WT and R124A 1 ps difference spectra. These spectra show the same overall spectral structure, with the most noteworthy differences being shifts in peak positions from 1577 to 1573 cm⁻¹, and from 1637 to 1632 cm⁻¹.

Figure S5.12. Photodiode array spectra of R124A vs. 0.2 mM NO at time points between 1 - 500 ms post-mixing. This shows a decrease in amplitude in the 424/432 nm reduced peaks and an increase at 396 nm (NO-bound). The respective decrease and increase in signal at 432 and 396 are illustrated inset.
Figure S5.13. NO concentration-dependence curve for R124A vs. NO acquired by SF/UV-vis. The rates were calculated by fitting of the 432 nm trace at various NO concentrations.

Figure S5.14. $k_1$ R124A laser-flash photolysis NO concentration-dependence curve. The plot shows a linear dependence, with a calculated $k_{on}$ value of $(3.50 \pm 0.15) \times 10^6 \text{ M}^{-1}\text{s}^{-1}$. 
Figure S5.15. $k_2$ R124A laser-flash photolysis NO concentration dependence curve. The plot shows a linear dependence, with a calculated $k_{on}$ value of $(7.23 \pm 0.92) \times 10^4 \, \text{M}^{-1}\text{s}^{-1}$. The 2 mM sample returned a lower rate than expected from the linear regression, which could be due to the different sample preparation method used for this sample (see Materials and Methods section).

Agreement of ΔA values between R124A TA and flash-photolysis experiments

The concentration of the R124A samples used for the TA and laser flash photolysis experiments were ~ 44 µM and ~ 9 µM respectively. These were set so that the Soret band (395 nm) absorption was 0.7 in the probe beam path (0.2 and 1 cm, respectively). The cuvette pathlength for the pump beam differed between the systems (2 mm and 4 mm, respectively), so the absorption at the Soret band in the pump beam was 0.088 and 0.035, respectively. The transient absorption experiments were performed with a power density of ~ 5.7 mJ/cm$^2$ (1 µJ, 150 µm beam diameter), and the flash photolysis experiments with a power density of ~ 127 mJ/cm$^2$ (100 mJ, 1 cm beam diameter). Between 1 and 10 µs the change in absorbance values were ~ 0.011 and ~ 0.041 for the TA and the flash photolysis experiments, respectively.

After correction for the different absorption cross sections in the two experiments, the TA and laser-flash photolysis values for ΔA/A are: 0.13 and 1.17, respectively. If this value is then divided by the power density, the resulting values are 0.022 and 0.009. In a perfect system these values should be the same, however factors such as pump/probe beam overlap, and the differing spectral bandwidth of the pump beams (~ 40 nm and > 1 nm, respectively) will affect the values, so we feel that the values are exhibit enough similarity to show reasonable correlation between the systems.
6. Conclusions and Outlook

The overriding aim of this PhD was to determine the potential for studying protein systems, particularly tetrapyrrrole-based proteins, using a range of UV-vis and IR time-resolved techniques. It was hypothesised that UV-vis techniques can provide the opportunity to monitor reaction chemistry, whilst IR techniques would probe associated protein dynamics. To test this potential, two protein systems were selected for analysis; ethanolamine ammonia lyase (EAL) and cytochrome c’ (Cytcp). The major findings for each, and overall implications for future research, will be covered in turn. Furthermore, this PhD involved the characterisation and development of a stopped-flow FT-IR (SF/FT-IR) instrument in collaboration with TgK Scientific, and will also be summarised.

6.1 Stopped-flow FT-IR method development

An association with TgK Scientific, the technical capabilities of the SF/FT-IR instrument were characterised using test chemistry. The ester hydrolysis of methyl chloroacetate was selected for analysis since the reaction chemistry results in a number of IR spectral changes over 100’s ms (225, 226). The chemicals required for analysis were readily available and affordable, thus allowing for multiple experiments to be implemented under various instrumental settings, without the need for reagent economy. The capability of the instrument to generate high quality datasets, including difference spectra (relative to an early time point) and reaction kinetics were demonstrated. Furthermore, since the spectral resolution directly affects the temporal resolution of the instrument, the reaction was measured using various spectral resolutions and its relationship with the acquisition rate determined. Finally, using a range of experimental parameters, the optimal shot volume was established as 100 µl. This volume setting allows for the expulsion of spent solution from a previous experiment, whilst preserving reagent economy.

These data provided useful insights into the capabilities of the SF/FT-IR prior to study of proteins, which represent a more expensive (and finite) sample. Furthermore, following this characterisation of the instrument, I assisted TgK Scientific on business trips to universities with SF/FT-IR instruments in South Africa, USA and Germany, to demonstrate best practice when using the technique, particularly for analysis of proteins.

6.2 Ethanolamine Ammonia Lyase

For our study of EAL, initially free cobalamin in solution were analysed using ultrafast time-resolved infrared (TRIR) spectroscopy, which included the physiological coenzyme of EAL, AdoCbl. As predicted from their sophisticated structure, each of these cobalamin molecules returned rich spectroscopic IR features after photoexcitation, with
varying decay kinetics. The calculated rates of these TRIR findings closely resembled those previously reported by ultrafast UV-vis transient absorption (TA) spectroscopy. The valuable insight yielded from these initial free cobalamin experiments were the vibronic spectral features returned after photoexcitation. Computational DFT calculations of these TRIR difference spectra revealed that the majority of the transient signals report on vibrational excitation of the corrin, with each band representing coupled motions throughout the ring system.

The AdoCbl TRIR findings had a direct relevance to the EAL study presented in Chapter 4, since the EAL-AdoCbl equivalent experiment could be performed, thus allowing additional vibrational features arising from protein-AdoCbl contacts to be determined. Small but significant additional signals were observed in the amide I region of the spectrum, comprising a bleach at 1650 cm$^{-1}$ and corresponding transient at 1661 cm$^{-1}$. This suggests vibrational coupling between the cofactor and protein following photoexcitation, perhaps facilitated by Glu$^{287}$. SF/FT-IR experiments of EAL mixed with a range of substrates produced more substantial signals that coincided exactly with the additional TRIR signals. Isotopic labelling of the protein scaffold indicated that these signals arise from vibrational changes in the protein structure, representing the first real-time monitoring of a protein dynamic contribution during AdoCbl-dependent catalysis. Furthermore, the calculated rates from these SF/FT-IR studies matched those of stopped-flow UV-vis (SF/UV-vis) experiments, indicating a coupling between protein dynamics and Co-C bond cleavage.

Previous publications have hypothesised that a mobile Glu$^{287}$ active site residue contributes to Co-C bond cleavage, substrate coordination, and subsequent radical stabilisation (101, 162, 163). From analysis of the SF/FT-IR data presented, the 1650 cm$^{-1}$ bleach represents a change in approximately 1 % of the overall IR absorbance at this wavenumber, which equates to 7 ± 2 peptide bond contributions. This finding is compelling, since it implicates, as predicted (162, 163), a relatively small protein structural change upon substrate binding, but one more substantial than a single glutamate residue illustrated in the substrate-bound crystal structure (101). This finding is perhaps unsurprising, and it remains likely that Glu$^{287}$ plays a major role in facilitating catalysis in EAL. However, for future research perhaps additional active site residues should receive attention, such as Asp$^{362}$, Gln$^{162}$, Arg$^{160}$, and Asn$^{193}$, all of which form hydrogen bond contacts with the substrate according to the crystal structure (101). Arg$^{160}$ in particular, has been shown previously to be critical for catalytic turnover (155), and research with regards to this residue could be revived.

Since the publication of the article presented in Chapter 4, studies to elucidate the importance of Glu$^{287}$ have been on-going in our group, with a view to extending the
mechanistic understanding of EAL catalysis. A range of EAL Glu\textsuperscript{287} variants have been generated, each of which affect the steric or electrostatic interactions with the substrate and 5'-deoxyadenosyl moiety. A glutamate to aspartate variant (E287D) retains the physiological negative charge, but removes the extension of the residue by one methyl group. A glutamine variant (E287Q) retains the steric influence of the residue, but removes the physiological negative charge of glutamate. Finally, an alanine variant (E287A) totally removes the mechanistic influence of the glutamate residue. Static UV-vis and SF/UV-vis analysis have suggested that this residue is indeed critical to both catalysis and protection from cobalamin inactivation under aerobic conditions (Chen, Zietek, Russell et al., in press). These studies have involved steady-state assays of the EAL variants and rapid-mixing of EAL with substrates, with vibrational and photoexcitation studies also planned. Study of these variants using TA and TRIR spectroscopy may allow us to determine whether previously observed protein dynamic signals following Co-C bond photolysis were due to Glu\textsuperscript{287} interacting with the 5'-deoxyadenosyl, thereby implicating its role in radical pair generation and stabilisation. Similarly, SF/FT-IR measurements can be coupled to our reported SF/UV-vis studies, to establish the overall importance of the Glu\textsuperscript{287} residue during catalysis, or whether other active site residues are sufficient to guide catalysis.

6.2.1 Implications for other AdoCbl-dependent enzymes

All AdoCbl-dependent enzymes undergo AdoCbl Co-C bond homolysis upon interaction with substrate yielding a radical pair that mediates catalysis. However, enzymes have been shown to undertake various structural dynamics in order to prompt Co-C homolysis (284). Unlike EAL, some AdoCbl-dependent enzymes adopt large scale domain motions in order to achieve active site priming upon substrate binding. Ornithine 4,5-aminomutase is one such enzyme, by adopting an open conformation for the substrate-free structure, with the AdoCbl binding domain and substrate binding domain separated by a substantial distance (258). Computational studies suggest that upon substrate binding ornithine 4,5-aminomutase assumes a closed state, with electrostatic and steric contacts with the AdoCbl facilitating Co-C bond cleavage (259). Methylmalonyl CoA mutase also undergoes large domain motions upon substrate binding, similarly converting from an open to closed conformation to generate radical pair formation (96, 285, 286). However, computational studies have indicated that in addition to these large domain motions upon substrate binding to methylmalonyl CoA mutase there is also a reorganisation of a number of active site residues (287). This suggests that, although on the surface AdoCbl-dependent enzymes may appear to adopt different modes for radical pair generation, they in fact employ similar mechanisms of catalysis. Indeed, in addition to methylmalonyl CoA mutase, computational
studies of glutamate mutase also suggest stabilisation of the 5'-deoxyadenosyl radical following Co-C cleavage in a manner analogous to EAL \((288)\).

Considering these apparent similarities between the various AdoCbl-dependent isomerases, this adds a broader relevance to our EAL vibrational spectroscopy findings. In all cases it appears that a negatively charged active site residue (normally glutamate) is heavily implicated in guiding catalysis. Therefore it is likely that our conclusions stating the coupled nature of protein dynamics to Co-C bond homolysis stands true for the majority of AdoCbl-dependent isomerases. Furthermore, future studies using Glu\(^{287}\) variants could also act as an indicator for AdoCbl-dependent enzyme catalysis more generally. In conclusion, the understanding of this curious group of enzymes, which harness the catalytic power of the large coenzyme AdoCbl, has been assisted by recent publications of crystal structures, and the dynamic contribution of the protein scaffold in guiding catalysis appears to be revealing itself.

### 6.3 Cytochrome c’

Our study of Cytcp focussed on how specific haem pocket residues contribute to the reactivity of the haem towards its physiological ligand (NO), and protection against solvent escape. Here, samples of various Cytcp-NO complexes were analysed using photoexcitation techniques, which involved the photodissociation of NO from haem, followed by probing of electronic and vibrational chemistry upon rebinding. UV-vis photoexcitation techniques (TA and flash-photolysis) allowed the probing of spectroscopic signals from fs – s following excitation, whilst TRIR probed vibrational geminate recombination events from fs – ns. In order to characterise physiological spectra, initially a wild-type AxCytcp (WT)-NO sample was probed. Following these experiments, the dynamic contributions of proximal arginine and distal leucine haem binding pocket residues were analysed, using two alanine variants (R124A and L16A, respectively).

The WT returned TA and flash photolysis results similar to those published previously \((222, 223)\), and TRIR allowed the direct probing of the 5c Fe-NO bond during geminate recombination. These findings suggested a tightly regulated process following NO photodissociation, with the majority (> 95 %) of NO undergoing prompt geminate recombination. A small proportion of WT sample allowed NO solvent escape and proximal histidine rebinding, which was illustrated by the minor signal remaining after ∼ 100 ps in the TA experiment. L16A binds NO with 6-coordinate geometry on the distal face \((218)\), therefore after photolysis this variant exhibited full geminate recombination in both the TA and TRIR within ∼ 1 ns. R124A returned similar TRIR difference spectra to WT, as a result of the 5c-NO binding of this variant, but illustrated significant differences in its UV-vis spectroscopic
features. Here, enhanced spectral complexity was observed for TA difference spectra most likely owing to the mixture of NO proximal and distal binding in the R124A sample. Furthermore, these signals demonstrated an extension of signal change into the μs time window and an increased residual difference signal after 10 μs. Together, these findings suggested a loss of the tight haem regulation demonstrated for WT, including an increase in NO solvent escape. Based on the time scales at which various protein motions / chemistry occur (covered in Section 1.2.4), an NO/His rebinding mechanism immediately following photolysis was proposed, with signals from fs – ns reporting on NO geminate recombination and His rebinding, and longer ns – μs signals constituting domain motions and haem readjustment. Finally, laser-flash photolysis of R124A in complex with NO revealed an additional NO concentration-dependent step, which was previously unresolved by SF/UV-vis due to the dead-time limitations of the instrument. This “fast” solvent rebinding step was assigned to formation of the distally bound 5c-NO, followed by "slow" proximal binding 5c-NO.

Overall, this study of Cytcp emphasised the importance of Arg124 in facilitating haem reactivity towards NO, and protection against solvent escape. This residue had previously been implicated in stabilising the haem and extending the lifetime of a 6c-NO intermediate during NO binding (220). It was shown in the same publication by SF/UV-vis that as the nature of this arginine is adjusted from the physiological form, the lifetime of the 6c His-Fe-NO intermediate is diminished (220). Thus, a valuable future experiment for Cytcp would be to repeat these photoexcitation techniques with an R124 variant more closely related to the WT. For instance, R124Q, where arginine is replaced with glutamine, presents a much diminished 6c His-Fe-NO intermediate during NO binding, but still retains 100 % proximal 5c-NO final product (220). This sample would act as a useful comparative study for the R124A TA spectroscopy data, which was complicated by the mixture of proximally and distally bound 5c-NO. This would allow the exact quantification of solvent escape, as well as adding support to the hypothesised R124A NO/His rebinding mechanism.

### 6.3.1 Implications for soluble guanylate cyclase

Soluble guanylate cyclase (sGC), the crucial NO-sensory protein in higher organisms, shares a number of mechanistic and ligand-binding properties with Cytcp. As a result of the inaccessibility to a 3D crystal structure of sGC, greater understanding of Cytcp serves as a homologous system for unravelling sGC function. The Cytcp findings presented in this thesis have emphasised the importance of Arg124 in facilitating haem reactivity and protection against solvent escape, to go with previous findings suggesting its role in extending the lifetime of the 6c His-Fe-NO intermediate (220). Interestingly, within the sGC β1 N-terminal,
which is proposed to contain the haem-binding pocket (289), there is also a conserved arginine residue at position 139. In the absence of structural information demonstrating the proximity of Arg^139 to the haem, claims of its importance in sGC are merely hypothetical. However, considering the similarly in the ligand binding mechanisms of Cytcp and sGC it is conceivable that this Arg^139 of sGC has a similar function to that of Arg^124 of Cytcp. With this in mind, it has been previously suggested that triggering of sGC catalytic activity could be directly mediated by the displacement of the proximally-ligated histidine residue by NO (211). Arg^124 has been shown by our data and previous studies (220) to directly influence the fate of this proximal histidine residue, since it ensures 100% proximal NO binding and stabilisation of the haem during NO binding (220), and increases the reactivity of the haem towards NO, thus preventing NO solvent escape and histidine rebinding. As such, in sGC, the related Arg^139 residue could be crucial in maintaining displacement of the histidine residue and therefore the continued catalytic activity of sGC. Such mechanistic insights could be crucial in determining the complex structural and chemical properties of this therapeutically-implicated enzyme.

6.4 Final comments and future applications

In studying EAL and Cytcp, the potential of these wide time scale spectroscopic techniques for the study of tetrapyrrole-based protein systems has been demonstrated. In terms of the photoexcitation techniques, these have exploited the light absorbing properties of tetrapyrroles to elicit electronic and vibrational changes in the cofactor and associated protein structure. In both cases, absorption of light results in photodissociation of a bound axial ligand. For EAL, this concerns an axially bound 5’-deoxyadenosyl group that forms part of the physiological cofactor, whereas with Cytcp, an exogenous NO ligand is photodissociated. EAL demonstrates how thermally-induced structural changes observed during catalysis can be mimicked by light excitation and therefore protein-cofactor contacts analysed. For Cytcp, the protein-ligand contacts are the focus of study, which can provide insights on the roles of important haem binding pocket residues. Essentially, these photoexcitation findings demonstrate the versatility of these techniques when analysing light-absorbing systems. Photosensory proteins utilise light as part of their function, and therefore are suitable for study by TA, TRIR and flash photolysis, amongst others. Here we have shown that protein dynamics can be probed in non-photosensory systems, exploiting the ability to photolyse chemical bonds and observe their recombination over a broad range of time scales. By way of amino acid substitutions, the contribution of individual protein residues contained in the cofactor binding site can be identified, thus allowing mechanistic insight into protein function.
As shown in our study of EAL, SF/FT-IR coupled with SF/UV-vis provides the opportunity to couple the oxidation state of the chromophore to related protein dynamics upon mixing of enzyme and substrate, or protein and ligand. Colour change upon initiation of function is a common trait of tetrapyrrole-based proteins, and as such SF/UV-vis has been used for a number of years. However, with SF/FT-IR, this allows signals arising from the protein scaffold to be identified, which may be otherwise spectroscopically silent in the UV-vis. Our studies have shown that isotopic labelling of proteins can be used to identify protein dynamic contributions during reaction, which can also be probed by amino acid substitution.

Finally, the advantage of coupling UV-vis and IR techniques has been emphasised, by complementing the results of one another and yielding synergistic insights into protein dynamics during function.
7. References


the influence of solvent on the primary photolysis mechanism and geminate recombination of methyl B12 coenzymes: comparison of the primary photolysis mechanism in methyl-, ethyl-, n-propyl-, and 5'-deoxyadenosylcobalamin, J. Am. Chem. Soc. 124, 434-441.


and a3 Redox Reactions in Cytochrome c Oxidase from Paracoccus denitrificans: Separation of Heme a and a3 Contributions and Assignment of Vibrational Modes, Biochemistry 38, 1685-1694.


8. Appendix

8.1 Ethanolamine Ammonia Lyase Overexpression / Purification Protocol

The EAL sample utilised for analysis was overexpressed and purified using the following protocol. Information on proof of successful yield is also provided.

8.1.1 Protocol

**Growth Medium**

For all steps in EAL preparation, 2-YT growth medium is used that contains:

- 16 gdm⁻³ tryptone
- 10 gdm⁻³ yeast extract
- 5 gdm⁻³ NaCl

Add 1 litre ddH₂O to every 31 g of media powder, autoclave at 2 bar, 121 °C for approx. 20 minutes. Following cooling, add stock ampicillin (100mgml⁻¹) to render final concentration 100 µgml⁻¹.

**Transformation into Competent Cells**

The plasmid ‘pET SEAL’ contains genes encoding both the 31 kDa β subunit and 50 kDa α subunit of EAL, as well as an ampicillin-resistance gene. The EAL subunit genes are under the control of the promoter for IPTG. Under sterile conditions, transform *E. coli* BL21*DE3* competent cells using the following protocol:

- Suspend a small quantity of thawed, competent cells in 100 µl TFB transformation buffer (10 mM K-MES, 100 mM KCl, 45 mM MnCl₂.4H₂O, 10 mM CaCl₂.2H₂O, and 3 mM HACoCl₃, pH 6.2, sterile filtered).
- To 50 µl of this competent cell solution, add 0.
- 2 µl of pET SEAL (concentration 100 ng µl⁻¹, therefore 20 ng of plasmid added) and incubate for 30 minutes on ice.
- Heat shock at 42 °C for 20 seconds, followed by 2 minutes on ice.
- Add 250 µl SOC recovery medium, then incubate for 1 hour at 37 °C to allow the repair of the cell membrane.

---

**EAL Overexpression**

**Agar Plates**
- Prepare LB agar plates under sterile conditions by adding 40 ml agar medium and 40 µl stock ampicillin to a falcon tube.
- After cooling, add 100 µl of the transformant mixture per plate then evenly spread across the surface with a sterile spreader bar.
- Incubate the agar plates overnight at 37 °C.

**2 × 10 ml Starter Cultures**
- Add 10 ml of sterilised 2-YT broth to sterile 20 ml tubes followed by 10 µl of stock ampicillin.
- For each tube, pick an isolated colony from the overnight agar plates using a sterilised pipette tip and add to the media.
- Incubate at 37 °C for ~ 4 hours in stirring incubator at 220 rpm, or until growth is observed.

**2 × 100 ml Starter Cultures**
- Add 100 ml of 2-YT to 250 ml conical flasks then autoclave.
- After this sterile broth has cooled, add 100 µl of ampicillin stock solution.
- Add a 10 ml starter culture to each of the 250ml conical flasks and incubate overnight at 37 °C in stirring incubator at 220 rpm.

**12 × 900 ml Cultures**
- Add 900 ml of 2-YT to 2 litre conical flasks then autoclave.
- After this sterile broth has cooled, add 910 µl of ampicillin stock.
- Under sterile conditions, add 15 ml of the 100 ml starter cultures to each of the sterile 900ml 2-YT broth media and incubate at 37 °C in a stirring incubator at 220 rpm for 2 hours.

**Induction**
- Blank a spectrophotometer at 600 nm using 2-YT broth.
- After 2 hours of incubation, aseptically remove 1 ml aliquots from two separate flasks, then measure the OD of the culture.
- Once the OD lies between 0.8 - 1.0 a.u., the cultures are ready for induction.
- Remove all 12 flasks from the stirring incubator and under aseptic conditions, add 910 µl of IPTG stock (0.1 M).
- Reduce temperature to 25 °C and incubate overnight.
**Harvest of Cells**

- Recover the overnight culture from the stirring incubator and transfer to 1 litre centrifuge bottles.
- Balance these bottles and centrifuge at 6000 RPM for 20 minutes at 4 °C.
- Discard the supernatant by pouring away from the pellet. The supernatant should be transparent, whilst the pellet should appear a "wet sand" colour.
- Pre-weigh several 50 ml conical flasks then scrape pellet in ¾ of the way up.
- Weigh this pellet and record.
- Snap freeze in liquid N₂ then store at -80 °C.

**EAL Purification**

The purification method is based on the limited solubility of EAL near neutral pH in buffers that contain glycerol. All steps are carried out under temperature controlled conditions set to 4 °C.

**Preparation of required buffers:**

- 1 M HEPES buffer: 47.66 g HEPES in 200 ml ddH₂O
- 50 mM HEPES/glycerol 10 % (v/v) pH 7.5: 5 ml 1 M HEPES in 100 ml ddH₂O
- 50 mM HEPES pH 7.5: 20 ml 1 M HEPES in 400 ml ddH₂O

**Cell Lysis**

- Thaw cell pellet on ice.
- Resuspend pellet in 100 ml of 50 mM HEPES/glycerol (10%) buffer at pH 7.5.
- Add a final concentration of 1 mM magnesium chloride solution and Complete™ EDTA-free protease inhibitor tablets (1 required per 50ml of suspension).
- After full resuspension, add 1 mg DNAse and 50 mg lysozyme.
- Sonicate the suspension mixture (light brown in colour) for 20 seconds at 40-45% intensity, followed by a 1 minute relaxation period including swirling. Repeat 15 times.
- After sonication the suspension should have a dark brown cream colour.
- Balance and spin the suspension at 18,000 rpm for 1 hour.
- Take a small sample of both supernatant and pellet (S1 and P1, respectively) for SDS-PAGE.
- Retain the pellet and resuspend in 250 ml of 50 mM HEPES buffer at pH 7.5.
- Add protease inhibitor tablets as required.
- Balance and spin suspension at 18,000 rpm for 30 minutes.
- Retain a small sample of the supernatant and pellet (S2 and P2, respectively) for SDS-PAGE.

- Retain the supernatant, which is pale-yellow in colour.

**Ammonium Sulphate Cut**

- Gradually add final concentration 84 gL⁻¹ NH₄SO₄ to supernatant.
- Stir for 20 minutes.
- Balance and spin mixture at 18,000 rpm for 30 minutes.
- Retain a small sample of the supernatant and pellet (S3 and P3, respectively) for SDS-PAGE.

**1st Dialysis**

- Resuspend pellet in 100 ml of 10 mM HEPES buffer at pH 7.5.
- Dialyse overnight with stirring against 5 L of 10 mM HEPES buffer at pH 7.5.

**2nd Dialysis**

- Dialyse the solution against 4 L of 10 mM HEPES for 2 × 1 hour.
- Balance and spin at 18,000 rpm for 30 mins.
- Retain a small sample of the supernatant and pellet (S4 and P4, respectively) for SDS-PAGE.
- Retain the supernatant, which should contain EAL and discard the pellet.

**Concentration of EAL**

- Fill 4 x 100,000 MWCO 'Millipore' centrifugal-filter devices with EAL solution.
- Balance and spin at 30,000 rpm for ~ 30 minutes.
- Top up the tubes and re-balance until volume has reduced by approx. 50%.
- Retain a small sample of the flow through and concentrated solutions (S5 and S6, respectively) for SDS-PAGE.
- Aliquot the concentrated sample and snap freeze in liquid N₂ then store at -80 °C.

**8.1.2 Determination of sample purity**

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of all retained samples is implemented to monitor the efficiency of the purification protocol, and to check the purity of the final enzyme sample. In this instance SDS acts as the denaturing agent and β-mercaptoethanol (β-ME) is added during sample preparation to reduce protein disulphide bonds.
Table 8.1. SDS-PAGE sample preparation for each of the EAL purification fractions.

<table>
<thead>
<tr>
<th>Sample Code</th>
<th>Sample (µl)</th>
<th>Sample Buffer (µl)</th>
<th>β-ME (µl)</th>
<th>H₂O (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1-P3</td>
<td>2 mm² pellet</td>
<td>9</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>S1/S6</td>
<td>0.5</td>
<td>9</td>
<td>1</td>
<td>19.5</td>
</tr>
<tr>
<td>S2/S3</td>
<td>4</td>
<td>9</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>S4</td>
<td>2</td>
<td>9</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>S5</td>
<td>9</td>
<td>9</td>
<td>1</td>
<td>11</td>
</tr>
</tbody>
</table>

- Place samples in a 100 °C heat block for 5 minutes
- Centrifuge samples at 13,000 rpm for 1 minute.
- Pre-run 12 % acrylamide GeBa gel (ChemBio) for 12 minutes at 160 V and 45 mA.
- Remove the comb and add running buffer to the wells followed by addition of 10 µl of samples and a molecular weight marker.
- After a running time of 1 hour, stain/destain the gel using coomassie blue stain.

Figure 8.1 provides an example gel following SDS-PAGE of each of these fractions. The star symbols represent the location of EAL protein according to protocol, and α/β symbols represent the bands for each of the subunits of EAL at 49 kDa and 30 kDa, respectively.

Figure 8.1. Example gel from SDS-PAGE of EAL purification fractions. S6 constitutes the final, pure sample containing the large (α) and small (β) subunit bands.
8.2  Steady-state assay of EAL

This coupled steady-state assay allows the spectral monitoring of EAL turnover, therefore allowing the calculation of reaction rates. The assay is coupled to yeast alcohol dehydrogenase (YADH); an enzyme that uses NADH as an electron donor to catalyse the conversion of acetaldehyde to ethanol, producing NAD⁺ as a by-product (Equation 8.1). Since NADH gives a strong absorption signal in the UV-vis at 340 nm, its consumption during YADH catalysis can be monitored spectroscopically. EAL catalyses the conversion of ethanolamine (2AE) to acetaldehyde and ammonia (Equation 8.2), and therefore can be coupled to YADH catalysis since this reaction provides the substrate for YADH. Thus, the EAL rate of reaction can be determined by the consumption of NADH at 340 nm.

\[
\text{Acetaldehyde} + \text{NADH}^{\text{YADH}} \rightarrow \text{Ethanol} + \text{NAD}^+ \quad \text{(Eqn. 8.1)}
\]

\[
\text{Ethanolamine}^{\text{EAL}} \rightarrow \text{Acetaldehyde} + \text{Ammonia} \quad \text{(Eqn. 8.2)}
\]

Therefore this assay involves a reaction cell containing YADH, NADH, 2AE and EAL holoenzyme (to be added last) and the monitoring of absorbance changes at 340 nm, which is directly related to the activity of EAL.

8.2.1  Reagent Preparation

**Buffer:** Add 0.953 g of HEPES powder to 200 ml of H₂O then adjust pH to 7.5.

**YADH:** 100 units are desired in the final reaction vessel. Add 1 mg to 1 ml of buffer then prepare a 1 in 10 dilution sample for measurement at 280 nm in a spectrophotometer. Subject this absorbance value to the following calculations (14.6 represents the extinction coefficient ε¹% in water at 280 nm, which can be used to generate the mg ml⁻¹ value):

- **Calculation of mg ml⁻¹:** \( \frac{\text{(Absorption Value)}}{14.6} \times 100 = x \text{ mg ml}^{-1} \)
- **Calculation of units:** \( x \text{ mg ml}^{-1} \times 340 \text{ units mg}^{-1} = x \text{ units ml}^{-1} \)
- **Calculation of volume:** \( \frac{100 \text{ units}}{x \text{ units ml}^{-1}} \times 1000 = \text{Number of µl required} \)
**NADH:** The desired final concentration is 150 nM. Add 3.82 mg to 500 µl buffer then prepare a 1 in 166.7 dilution (6 µl in 1000 µl) sample for measurement at 340 nm in a spectrophotometer. Subject this absorbance value to the following calculations:

- Calculation of concentration: \( \frac{(\text{Absorption Value})}{6.22} \times 166.7 = x \mu M \)
- Calculation of volume: \( \frac{150 \text{ nM}}{(x \mu M \times 1000)} \times 1000 = \mu l \text{ required} \)

**2AE:** The desired final concentration is 1 mM, therefore add 0.84 µl to 7 ml buffer, then add 500 µl to the reaction vessel (1000 µl capacity).

**EAL:** The desired final concentration is 100 nM, therefore calculate the stock concentration and dilute accordingly. Before adding to the cuvette, the EAL apoenzyme should be incubated with ~ 10× excess of 5′deoxyadenosylcobalamin (AdoCbl) on ice for 10 minutes to ensure binding of the coenzyme.

### 8.2.2 Protocol

- After calculating the volumes of each of the reagents required, add all these reagents except EAL holoenzyme to the 1 ml cuvette and add buffer to make up to 1 ml (minus the volume of EAL to be added).
- Allow this to equilibrate to 25 °C in the cuvette holder of the UV-vis spectrophotometer.
- Meanwhile, using the “Kinetics” programme of the Cary® 50 spectrophotometer (Varian, Inc.), prepare a method that monitors the change in absorbance at 340 nm and set the measurement time to 1 minute.
- Following equilibration, add EAL and mix thoroughly and quickly then begin measurement of absorbance change.
- A steady decrease in absorbance signal should be observed, with the greatest rate of increase at the start of the reaction (the initial velocity).
- Measure the initial decrease in absorption (Absorbance/min) using the kinetics programme. This can be converted into the more conventional µM/sec by the following steps:
- Calculation of mM/min: \( \frac{(\text{Absorption/min})}{6.22} = \text{mM/min} \)
- Calculation of µM/min: \( \text{mM/min} \times 1000 = \text{µM/min} \)
- Conversion to µM/sec: \( \frac{(\text{µM/min})}{60} = \text{µM/sec} \)

### 8.2.3 Activity Assay Standard Curve

An activity assay standard curve was generated by assaying known concentrations of EAL between 5 – 100 nM then plotting against the initial rate. This standard curve can be used to estimate the enzyme activity of a sample during a purification protocol or otherwise. These data were generated using the above protocol, with three replicates generated at six different EAL concentrations. The resulting standard curve is illustrated in Figure 8.2.

![Figure 8.2. EAL steady-state assay standard curve. The relationship between initial rate and EAL concentration is presented, and the equation of the line of best fit detailed.](image)
8.3 **\(^{15}\)N-isotopically labelled EAL protocol**

The SF/FT-IR analysis of EAL presented in Chapter 4 required \(^{15}\)N-labelling of EAL apoprotein. As covered in the materials and methods, this employed a minimal medium method but overall had a similar overexpression and purification protocol to 2-YT broth medium. To prepare the minimal media the following formulation is utilised:

*Solution A:* (487.5 ml H\(_2\)O)

- \(\text{Na}_2\text{HPO}_4\) 7.3 g
- \(\text{KH}_2\text{PO}_4\) 2.7 g

Adjust to pH 7.2 then autoclave

*Solution B:* (12.5 ml H\(_2\)O)

- Glucose 2 g
- \(\text{MgSO}_4\) 120 mg
- \(^{15}\text{NH}_4\text{Cl}\) 0.5 g
- \(\text{MnSO}_4\) (5 mg ml\(^{-1}\)) 0.25 ml
- \(\text{CaCl}_2\) (37.5 mg ml\(^{-1}\)) 0.25 ml
- Thiamine chloride 10 mg

After adding thiamine, filter sterilise

Thiamine chloride is light and temperature sensitive when in solution, and therefore should be added last, prior to filter sterilising and inoculating with *E. coli* culture. Following cooling of autoclaved and sterilised media, combine the two solutions and add stock ampicillin (100 mg ml\(^{-1}\)) to render final concentration 100 µg ml\(^{-1}\).

8.3.1 **Trial of minimal medium protocol**

8.3.1.1 **EAL activity trial using minimal medium**

Initially, to ensure the minimal medium protocol yielded sufficiently active and pure EAL, a trial run was implemented that used non-isotopically labelled NH\(_4\)Cl. This trial was necessary due to the relative expense of \(^{15}\text{NH}_4\text{Cl}\) and the quantities of the chemical required. After transformation of *E. coli* strains with the EAL plasmid, a 10 ml minimal medium starter culture was inoculated with an isolated colony. This was incubated at 37 °C overnight then split, with half added to 500 ml minimal medium and half added to 500 ml 2-YT broth. These cultures were incubated at 37 °C until OD = 0.6, then induced using IPTG and the temperature reduced to 25 °C.

After overnight incubation these flasks were centrifuged and kept separate on ice. Both samples were resuspended in 50 ml of 50 mM HEPES pH 7.5 buffer containing protease.
inhibitor tablets and MgCl$_2$ (1 mM). Each sample was sonicated and centrifuged at 18,000 RPM for 1 hour. The supernatant from each of these samples was then assayed for activity using the activity assay detailed in Section 8.2. A 1 in 1,000 dilution sample of supernatant was used for the assay from both the minimal medium and 2-YT broth sample with the following results:

<table>
<thead>
<tr>
<th>Medium</th>
<th>Initial rate</th>
<th>Activity standard curve, equates to</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal medium</td>
<td>1.31 µM s$^{-1}$</td>
<td>54.8 µM EAL</td>
</tr>
<tr>
<td>2-YT broth</td>
<td>3.31 µM s$^{-1}$</td>
<td>139 µM EAL</td>
</tr>
</tbody>
</table>

As expected, when grown in the rich medium, a larger yield is observed. However, when using minimal media a significant activity is still observed and therefore expression was successful.

8.3.1.2 Determination of EAL yield and purity

In order to determine the volume required to achieve adequate EAL concentrations for SF/FT-IR analysis, a trial grow up was implemented using non-isotopically labelled NH$_4$Cl. To begin, 18 litres were used for overexpression, which yielded a cell pellet of 37 g. Since a pellet of approximately 90 g is normally purified to yield EAL stock of the desired concentration, this protocol was repeated, which yielded a pellet of 40 g. These cell pellets were then purified using the protocol outlined in Section 8.1.1. At each centrifugation stage a sample of the pellet and supernatant was taken for determination of purification success by SDS-PAGE. The resultant gel from this trial overexpression and purification is illustrated in Figure 8.4. This gel contains the same sample labels as used in Figure 8.1, and confirms the successful yield and purity of EAL using the minimal medium protocol. After sample concentration, approximately 20 ml of EAL was retained, of holoenzyme concentration 32 µM. This trial indicated that if 72 flasks of 0.5 L minimal medium were prepared and purified, an EAL sample of sufficient volume and concentration could be achieved for SF/FT-IR analysis.
8.3.2 \textsuperscript{15}N-labelled EAL overexpression and purification

Following the successful trials using non-isotopically labelled NH\textsubscript{4}Cl, the minimal medium overexpression and purification protocol was repeated using \textsuperscript{15}NH\textsubscript{4}Cl. After concentration this yielded 12 ml of 80 µl holoenzyme EAL, which was sufficient for SF/FT-IR work. SDS-PAGE indicated that pure sample was obtained, and had bands equivalent to a previous non-isotopically labelled experiment (\textsuperscript{14}N EAL well), as shown in Figure 8.4. Furthermore, using the EAL coupled steady-state assay it was shown that this sample contained active EAL.
8.3.2.1 Confirmation of successful $^{15}$N-labelling of EAL by mass spectrometry

In order to confirm unequivocally that this EAL sample contained isotopically labelled protein, the sample was analysed by mass spectrometry, specifically using the matrix assisted laser desorption/ionisation time of flight (MALDI-TOF) method. This method allows the mass to charge ratio ($m/z$) to be easily determined and therefore if the charge is known, the mass can be elucidated. Samples of non-isotopically labelled and $^{15}$N-labelled EAL were analysed using a trypsin digest of EAL from SDS-PAGE. The mass spectrometry spectra for $\alpha$ and $\beta$ subunits of each of these EAL samples are illustrated in Figures 8.5 and 8.6, respectively. Figure 8.5 illustrates the large $\alpha$ subunit mass spectrometry spectra for both the $^{14}$N EAL sample (green trace) and $^{15}$N EAL sample (pink trace). Each have the same overall profile, but in the case of the $^{15}$N plot all the peaks have been shifted to the right, suggesting an overall mass increase in the sample. Figure 8.6 illustrates the small $\beta$ beta subunit mass spectrometry spectra for the $^{14}$N EAL sample (blue trace) and the $^{15}$N EAL sample (red trace). As for the $\alpha$ subunit spectra, there is an overall shift in $m/z$ ratio to the right, qualitatively indicating an increase in mass.

![Figure 8.5. $\alpha$ subunit mass spectrometry spectra for $^{14}$N-EAL and $^{15}$N-EAL samples.](image)
Figure 8.6. β subunit mass spectrometry spectra for $^{14}$N-EAL and $^{15}$N-EAL samples.

Since the fragments were generated using a trypsin digest, the identities of the peptide fragments can be determined using the EAL amino acid sequence and the "Expasy PeptideMass" online tool. The peptide fragment masses could then be paired with the fragment masses returned from the $^{14}$N-EAL data. Although the m/z ratio is different between the $^{14}$N-EAL and $^{15}$N-EAL fragments, the "retention time" was identical, allowing the pairing of $^{14}$N and $^{15}$N fragments, to go with amino acid identity of the fragment. All amino acids contain one nitrogen atom in the peptide backbone, with some amino acids containing additional nitrogen atoms in their side chains, as detailed in Table 8.2. With this information, the increase in mass as a result of $^{15}$N substitution can be predicted. Selected α and β fragment data are detailed in Tables 8.3 and 8.4, respectively, which compare the peptide amino acid identities of fragments, and therefore the number of nitrogen atoms, with the difference in mass between $^{14}$N- and $^{15}$N-EAL fragments. Since the mass increase correlates with the number of nitrogen atoms in the individual fragments within error, the success of the isotopic labelling of EAL can be confirmed.
### Table 8.2. Amino acids containing N atoms in their variable side chain.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Single letter code</th>
<th>Additional side chain N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arginine</td>
<td>R</td>
<td>3</td>
</tr>
<tr>
<td>Asparagine</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Q</td>
<td>1</td>
</tr>
<tr>
<td>Histidine</td>
<td>H</td>
<td>2</td>
</tr>
<tr>
<td>Lysine</td>
<td>K</td>
<td>1</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>W</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table 8.3. Comparison of database peptide data with α-subunit mass spectrometry of $^{14}$N-EAL and $^{15}$N-EAL. The differences in mass between the heavier $^{15}$N-EAL fragments and lighter $^{14}$N-EAL correlate to the number of nitrogen atoms contained within the peptide, as confirmed with the amino acid identity.

<table>
<thead>
<tr>
<th>Database mass</th>
<th>Peptide identity</th>
<th>No. N Atoms</th>
<th>$^{14}$N-EAL mass</th>
<th>$^{15}$N-EAL mass</th>
<th>Mass Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>586.3585</td>
<td>MPVIK</td>
<td>6</td>
<td>586.3477</td>
<td>592.3335</td>
<td>5.9858</td>
</tr>
<tr>
<td>692.3678</td>
<td>NYGLAR</td>
<td>10</td>
<td>692.3561</td>
<td>702.3297</td>
<td>9.9736</td>
</tr>
<tr>
<td>852.4090</td>
<td>AGDPSLFF</td>
<td>8</td>
<td>852.7594</td>
<td>860.3766</td>
<td>7.6172</td>
</tr>
<tr>
<td>862.4370</td>
<td>AVGAEFNR</td>
<td>12</td>
<td>862.4255</td>
<td>874.3878</td>
<td>11.9623</td>
</tr>
<tr>
<td>1008.5101</td>
<td>RAGDPSSLFF</td>
<td>12</td>
<td>1008.497</td>
<td>1020.459</td>
<td>11.962</td>
</tr>
<tr>
<td>1143.6572</td>
<td>GLTSEVVAAVAK</td>
<td>13</td>
<td>1143.645</td>
<td>1156.601</td>
<td>12.956</td>
</tr>
<tr>
<td>1271.7449</td>
<td>KGLTSEVVAAVAK</td>
<td>15</td>
<td>1271.741</td>
<td>1286.69</td>
<td>14.949</td>
</tr>
<tr>
<td>1316.6837</td>
<td>TTLFGNVYYQFK</td>
<td>14</td>
<td>1316.674</td>
<td>1330.622</td>
<td>13.948</td>
</tr>
<tr>
<td>1516.7554</td>
<td>SGDVLAGVAAASSQER</td>
<td>20</td>
<td>1516.752</td>
<td>1536.679</td>
<td>19.927</td>
</tr>
<tr>
<td>1549.7445</td>
<td>LIQDDVNAYQKR</td>
<td>19</td>
<td>1549.742</td>
<td>1568.675</td>
<td>18.933</td>
</tr>
</tbody>
</table>

### Table 8.4. Comparison of database peptide data with β-subunit mass spectrometry of $^{14}$N-EAL and $^{15}$N-EAL. The differences in mass between the heavier $^{15}$N-EAL fragments and lighter $^{14}$N-EAL correlate to the number of nitrogen atoms contained within the peptide, as confirmed with the amino acid identity.

<table>
<thead>
<tr>
<th>Database mass</th>
<th>Peptide identity</th>
<th>No. N Atoms</th>
<th>$^{14}$N-EAL mass</th>
<th>$^{15}$N-EAL mass</th>
<th>Mass Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>700.4232</td>
<td>TQALLR</td>
<td>10</td>
<td>700.4192</td>
<td>710.3933</td>
<td>9.9741</td>
</tr>
<tr>
<td>844.4191</td>
<td>FLADHRSR</td>
<td>12</td>
<td>844.4116</td>
<td>856.3783</td>
<td>11.9667</td>
</tr>
<tr>
<td>848.4174</td>
<td>ASGINMTR</td>
<td>12</td>
<td>848.4103</td>
<td>860.3776</td>
<td>11.9673</td>
</tr>
<tr>
<td>884.5079</td>
<td>AQGLLEVR</td>
<td>12</td>
<td>884.5007</td>
<td>896.4676</td>
<td>11.9669</td>
</tr>
<tr>
<td>960.4876</td>
<td>VATTVEADR</td>
<td>12</td>
<td>960.4775</td>
<td>972.4452</td>
<td>11.9677</td>
</tr>
<tr>
<td>1055.5862</td>
<td>LSPEAIDALK</td>
<td>11</td>
<td>1055.576</td>
<td>1066.546</td>
<td>10.97</td>
</tr>
<tr>
<td>1404.7513</td>
<td>QAGLNVGTPFFVR</td>
<td>18</td>
<td>1404.741</td>
<td>1422.687</td>
<td>17.946</td>
</tr>
</tbody>
</table>
8.4 Generation of Cytochrome c’ L16A and R124A variants

The findings presented in Chapter 5 required the generation of L16A and R124A variants of *Alcaligenes xylosoxidans* cytochrome c’ (*AxCytcp*). The protocol, by which these variants were produced, and the sequencing data illustrating successful mutagenesis is presented herein.

8.4.1 Plasmid design

For overexpression of *AxCytcp*, *E. coli* cells are transformed with the plasmid pTCP-1; a pET 26b(+) vector that contains the genomic sequence for the protein. The DNA sequence encoding *AxCytcp* is detailed below, with the ATG start codon in green, TAA stop codon in red, CTG codon encoding leucine-16 (L16) and the CGC codon encoding arginine (R124) in blue. Also detailed is the amino acid sequence that this DNA encodes. It should be noted that physiologically, following expression, the protein has a leader sequence to allow translocation to the periplasm where it is cleaved. This explains L16 and R124 having amino acid numbers logically lower than expected, since the fully-formed protein has a glutamine residue as the first residue, not methionine.

**DNA sequence of wild-type AxCytcp**

ATG\_AAGAAGTTGTCCGCCGCGTTGGCCTGTATGACCGTCGGATCGCTGCTGGCGACCTCGGCCCAGGCGCAGTTCGCCAAGCCCG
\_AAGATGCCGTCAAGTACCGCCAGTCGGCGGCTGACGCTGAAGGGCCCTGCACGCTCGCTTCGCCCGCAATGCCGCGGCGGCGGCGGCGGCTGGACCGCCCGGCGGCGACGCGGCCAGCTTCACGCTGTCGGCGGCTGCCCTGGGCGGCTTTCGGTCCCG GCACGGAAGGCGGC\_GACGCCCGTCCCGAGATCTGGAGCGACGCGGCCAGCTTCAAGCAGAAGCAGCAGGGTTCCAGGACAACATCGTCAAGCTGTCGGCGGCTGCCGACGCCGGCGACCTGGACAAGCTGCGCGCCGCATTCGGCGACGTGGGCGCGAGCTGCAAGGCCTGCCACGACGCCTATCGC
\_AAGAAGAA\_G\_TAA

**Amino acid sequence of wild-type AxCytcp**

MKKLSTLAALACMTVGSLLATSQAQQFAKPEAVKYRQSAFLTMASHFGRMTPVVKQAYAPIAQQIAKANVEKLTSALPWAAGFPSTEGG DARPHEIWSDAASFKQKQAFQDNIVKLSAADADGLDKLRRAAFDVGASCKACHDAYRKKK

For L16A the CTG codon encoding L16 was converted to GCC, whereas for R124A the CGC codon encoding R124 was converted to GCC. Therefore, the primers selected for production of L16A and R124A variants of *AxCytcp* were as follows.

**L16A primers**

*Forward:* 5’-CCGCCAGTCGGCGCGCCACGCTGATGGGCGC-3’

*Reverse:* 5’-GGCCATCGCCTGGCAGCTGACTGGCGG-3’

**R124A primers**

*Forward:* 5’-CCAGCCTGCTGCCAAAGAGATAAGC-3’

*Reverse:* 5’-GCTTACTCTCTCTGGCATAAGGGCTG-3’
8.4.2 L16A and R124A mutagenesis

For mutagenesis, a PCR reaction was carried out for each of these variants using the primers detailed. This utilised the Phusion™ polymerase, along with dNTPs and appropriate buffer supplied with Phusion™ enzyme. The following formulation was required for each of the reagents to create a 500 µl “master mix” to be aliquoted into separate PCR tubes. The polymerase was added last to this reaction mixture to preserve enzymatic activity. Also, the nuclease-free H₂O was HPLC-grade water that had been autoclaved.

**Table 8.5. Reagent volumes for mutagenesis of AxCytcp to L16A and R124A.** Detailed in the “master mix” for aliquoting into PCR tubes

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Stock concentration</th>
<th>Final concentration</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer</td>
<td>5×</td>
<td>1×</td>
<td>100</td>
</tr>
<tr>
<td>dNTPs</td>
<td>10 mM</td>
<td>0.2 mM</td>
<td>10</td>
</tr>
<tr>
<td>Forward Primer</td>
<td>20 µM</td>
<td>0.5 µM</td>
<td>12.5</td>
</tr>
<tr>
<td>Reverse Primer</td>
<td>20 µM</td>
<td>0.5 µM</td>
<td>12.5</td>
</tr>
<tr>
<td>DMSO</td>
<td>100 %</td>
<td>3 %</td>
<td>15</td>
</tr>
<tr>
<td>Phusion polymerase</td>
<td>100 units</td>
<td>1 unit</td>
<td>5</td>
</tr>
<tr>
<td>Template plasmid</td>
<td>10 µg</td>
<td>200 ng</td>
<td>10</td>
</tr>
<tr>
<td>Nuclease-free H₂O</td>
<td>-</td>
<td>-</td>
<td>332.5</td>
</tr>
</tbody>
</table>

This master mix was distributed into separate PCR tubes in 50 µl aliquots and then placed in a PCR machine with a temperature gradient for the annealing step. Thus, each of the aliquots is subjected to various annealing temperatures. A programme was constructed with the following parameters.

- Initial Denaturation step: 94 °C for 2 mins

25 cycles of:
- Denaturation: 94 °C for 30 secs
- Annealing: 55 – 65 °C for 1 min
- Extension: 72 °C for 6 mins

- Final extension step: 72 °C for 10 mins

After conclusion of the PCR, the restriction enzyme DpnI is added to the sample to degrade any parental DNA. DpnI identifies methylated DNA in conjunction with the palindromic DNA sequence shown below. Bacteria methylate DNA as a defence mechanism against bacteriophage infection, and any unmethylated DNA is digested by restriction endonucleases and exonucleases. Since our parental DNA was amplified in *E. coli*, this DNA contains methylated nucleotides, unlike DNA generated from PCR.
**DpnI Recognition Site:**

\[
\begin{align*}
5' & \quad G & A & T & C & \ldots & \ldots & 3' \\
3' & \quad C & T & A & G & \ldots & \ldots & 5'
\end{align*}
\]

Since the sample volume is 50 µl for each PCR product, 5 µl of 10× NEBuffer 4 and 2 µl of DpnI restriction enzyme (New England BioLabs, Inc.) of activity 20,000 units/ml (~ 700 units of activity post-mixing) were added to the PCR product. This was mixed thoroughly then incubated at 37 °C for 2 hours. The sample was then heat inactivated at 80 °C for 20 minutes, which denatures the restriction enzyme but not the DNA double helix.

### 8.4.3 Sequence data demonstrating successful mutagenesis

In order to confirm the successful generation of L16A and R124A variants, the plasmids produced by PCR were sequenced and compared to the genomic DNA and amino acid sequences. These results are illustrated below, with a “*” icon denoting a match between wild-type and variant DNA or amino acid. These sequence alignment data confirm successful mutagenesis of both L16A and R124A.

**DNA sequence alignment of wild-type and L16A AxCytcp**

```plaintext
wild-type  ---------------------------------------
           ATG
L16A       TAGAAAATTTTGTCTTAATTTAAAGGATCTATCAATAT 21
           ATG
wild-type  GCCGGCGCAGTTGCCGTGATGGACTGGCAGACTGCCAGGAGGCATCT 81
L16A       GCCGGCGCAGTTGCCGTGATGGACTGGCAGACTGCCAGGAGGCATCT 120
wild-type  GCCAAGCCCGAATGCGCTCGACCCGAGCTGCCAGGAGGCATCT 141
L16A       GCCAAGCCCGAATGCGCTCGACCCGAGCTGCCAGGAGGCATCT 180
wild-type  TCGGCCGCTAGACCGCAGTTGCCGTGATGGACTGGCAGACTGCCAGGAGGCATCT 201
L16A       TCGGCCGCTAGACCGCAGTTGCCGTGATGGACTGGCAGACTGCCAGGAGGCATCT 240
wild-type  GCCAACGTCAGATGGCTCAATACAAGCTCGACCTGCCAGGAGGCATCT 261
L16A       GCCAACGTCAGATGGCTCAATACAAGCTCGACCTGCCAGGAGGCATCT 300
wild-type  ACGGAAGCCGCGCAAGCCTGCCAGGAGGCATCT 321
L16A       ACGGAAGCCGCGCAAGCCTGCCAGGAGGCATCT 360
wild-type  CAGCAAGCGCCGCTACCTGCAATGGAAGGAGGCATCT 381
L16A       CAGCAAGCGCCGCTACCTGCAATGGAAGGAGGCATCT 420
wild-type  GACAGCCAGCGCCGCTACCTGCAATGGAAGGAGGCATCT 441
L16A       GACAGCCAGCGCCGCTACCTGCAATGGAAGGAGGCATCT 480
wild-type  TATCGCAAGAAGAAGTAAAGCCGTACCGGCTTGGCTTCGGCCTGGGCGCTTGGTCCCCGCC 486
L16A       TATCGCAAGAAGAAGTAAAGCCGTACCGGCTTGGCTTCGGCCTGGGCGCTTGGTCCCCGCC 540

wild-type  GACAAGCGCCGCTACCTGCAATGGAAGGAGGCATCT 486
L16A       GACAAGCGCCGCTACCTGCAATGGAAGGAGGCATCT 540
```
### Amino acid sequence alignment of wild-type and L16A AxCytcp

**wild-type**

KRLSTLAAALACMTVGSLLATSQAQFKPEDAVKYRQSATLMSH 47
L16A

NNFV-L-EGDIHRKKLSTLAAALACMTVGSLLATSQAQFKPEDAVKYRQSATLMSH 57

**wild-type**

FGRMTPVVKQQAQFYDAIQAKNEVKLTLALPWAASFPGETGSDFDAREPIWSDAASFKQK 107
L16A

FGRMTPVVKQQAQFYDAIQAKNEVKLTLALPWAASFPGETGSDFDAREPIWSDAASFKQK 117

**wild-type**

QAPQDNIVKL5AAADADLDKLRAAFDVAGSCKACHDAYKKK------------------ 152
L16A

QAPQDNIVKL5AAADADLDKLRAAFDVAGSCKACHDAYKKK-AGTAPCSEFELRR 176

### DNA sequence alignment of wild-type and R124A AxCytcp

**wild-type**

ATG AAGAAGTTGTCCACGCTCG 22
R124A

AGAATAATTTTTGTTAATCTTAAAGAAGGAGATATACAT 60

**wild-type**

CCGCGTTGGCCTGTATGACCGTCGGATCGCTGCTGGCGACCTCGGCCCAGGCGCAGTTCG 82
R124A

CCGCGTTGGCCTGTATGACCGTCGGATCGCTGCTGGCGACCTCGGCCCAGGCGCAGTTCG 120

**wild-type**

CCGCGTTGGCCTGTATGACCGTCGGATCGCTGCTGGCGACCTCGGCCCAGGCGCAGTTCG 142
R124A

CCGCGTTGGCCTGTATGACCGTCGGATCGCTGCTGGCGACCTCGGCCCAGGCGCAGTTCG 180

**wild-type**

TCGGCGCGACCGCGCCCTGGCTCTGGCGGTGCTCCGCGACTGTC 202
R124A

TCGGCGCGACCGCGCCCTGGCTCTGGCGGTGCTCCGCGACTGTC 240

**wild-type**

CCGCGCGACCGCGCCCTGGCTCTGGCGGTGCTCCGCGACTGTC 262
R124A

CCGCGCGACCGCGCCCTGGCTCTGGCGGTGCTCCGCGACTGTC 300

**wild-type**

CCGCGCGACCGCGCCCTGGCTCTGGCGGTGCTCCGCGACTGTC 322
R124A

CCGCGCGACCGCGCCCTGGCTCTGGCGGTGCTCCGCGACTGTC 360

**wild-type**

AGCAGGGGTTCGGCAAGCAACATCGTCAAGCTGTCGGCGGCTGCCGACGCCACCGCTGG 382
R124A

AGCAGGGGTTCGGCAAGCAACATCGTCAAGCTGTCGGCGGCTGCCGACGCCACCGCTGG 420

**wild-type**

ACAAGCTGCGCGCCGCTGGCAAGCAACATCGTCAAGCTGTCGGCGGCTGCCGACGCCACCGCTGG 442
R124A

ACAAGCTGCGCGCCGCTGGCAAGCAACATCGTCAAGCTGTCGGCGGCTGCCGACGCCACCGCTGG 480

**wild-type**

ATGCCAAGAAGAAGTACCGCGGCTGGCAAGCAACATCGTCAAGCTGTCGGCGGCTGCCGACGCCACCGCTGG 502
R124A

ATGCCAAGAAGAAGTACCGCGGCTGGCAAGCAACATCGTCAAGCTGTCGGCGGCTGCCGACGCCACCGCTGG 532

** Amino acid sequence alignment of wild-type and R124A AxCytcp**

**wild-type**

KRLSTLAAALACMTVGSLLATSQAQFKPEDAVKYRQSATLMSH 48
R124A

NNFV-L-EGDIHRKKLSTLAAALACMTVGSLLATSQAQFKPEDAVKYRQSATLMSH 58

**wild-type**

FGRMTPVVKQQAQFYDAIQAKNEVKLTLALPWAASFPGETGSDFDAREPIWSDAASFKQKQ 108
R124A

FGRMTPVVKQQAQFYDAIQAKNEVKLTLALPWAASFPGETGSDFDAREPIWSDAASFKQKQ 118

**wild-type**

QAPQDNIVKL5AAADADLDKLRAAFDVAGSCKACHDAYKKK------------------ 152
R124A

QAPQDNIVKL5AAADADLDKLRAAFDVAGSCKACHDAYKKK-AGTAPCSEFELRR 177

182
8.5 Cytochrome c’ Overexpression / Purification Protocol

The wild-type, L16A and R124A AxCytcp proteins analysed in the article presented in Chapter 5 were overexpressed and purified using the same common protocol. The AxCytcp gene is contained on the pTCP-1 plasmid (containing specific genes for L16A and R124A as outlined in Section 8.4), and a separate plasmid, “pEC86” carries the ccmA-H genes from *E. coli* necessary for the maturation of c-type cytochromes. Thus, BL21(DE3) *E. coli* cells must be transformed with both these plasmids prior to overexpression. pTCP-1 contains a gene encoding kanamycin resistance, and pEC86 contains a gene encoding chloramphenicol resistance.

8.5.1 Protocol

In all cases where LB medium or agar contain antibiotics, this is at a final concentration of 30 µgml\(^{-1}\) for kanamycin and 20 µgml\(^{-1}\) for chloramphenicol

**Generation of competent cells and transformation**

- Under sterile conditions, streak out BL21(DE3) cells onto LB agar plates.
- Incubate overnight at 37 °C.
- Inoculate 5 ml of sterile LB broth with an isolated colony from overnight agar plates and incubate overnight at 37 °C.
- Inoculate a fresh, sterile, 5 ml LB broth solution with 100 µl of the overnight culture and grow until the OD ~ 0.5.
- Transfer a 1 ml aliquot to a sterile microfuge tube and centrifuge at 13,000 rpm for 30 seconds.
- Resuspend in 0.5 ml of sterile, pre-chilled 50 mM CaCl\(_2\).
- Centrifuge at 13,000 rpm for 30 seconds.
- Resuspend in 0.5 ml of sterile, pre-chilled 50 mM CaCl\(_2\), and leave on ice for 1 hour.
- Add 1 µl of pTCP-1 (of concentration 100 ng µl\(^{-1}\), therefore 100 ng of plasmid added) to the competent cell sample.
- Leave on ice for 30 minutes.
- Heat shock at 42 °C for 30 seconds.
- Return sample to ice for 2 minutes.
- Incubate at 37 °C for 1 hour.

---

Transfer 100 µl of this transformant mixture to LB with kanamycin (LB/kan) agar plates and incubate at 37 °C overnight.

- Pick an isolated colony from the overnight agar plate and inoculate 5 ml LB/kan and incubate 37 °C overnight.
- Inoculate a fresh, sterile, 5 ml LB/kan broth solution with 100 µl of the overnight culture and grow until the OD ~ 0.5.
- Transfer a 1 ml aliquot to a sterile microfuge tube and centrifuge at 13,000 rpm for 30 seconds.
- Resuspend in 0.5 ml of sterile, pre-chilled 50 mM CaCl$_2$.
- Centrifuge at 13,000 rpm for 30 seconds.
- Resuspend in 0.5 ml of sterile, pre-chilled 50 mM CaCl$_2$, and leave on ice for 1 hour.
- Add 1 µl of pEC86 (of concentration 100 ng µl$^{-1}$, therefore 100 ng of plasmid added) to the competent cell sample.
- Leave on ice for 30 minutes.
- Heat shock at 42 °C for 30 seconds.
- Return sample to ice for 2 minutes.
- Incubate at 37 °C for 1 hour.
- Transfer 100 µl of this transformant mixture to LB with kanamycin and chloramphenicol (LB/kan/chlor) agar plates and incubate at 37 °C overnight.

**Overexpression**

- Pick an isolated colony from the overnight plate and inoculate a sterile, 5 ml LB/kan/chlor broth.
- Incubate for 6 hours at 37 °C.
- Transfer this solution to sterile, 500 ml LB/kan/chlor broth.
- Incubate overnight at 37 °C.
- Transfer 10× 50 ml aliquots of overnight culture to sterile solutions of 950 ml LB/kan/chlor.
- Incubate at 37 °C for 5 hours.
- Prepare a 0.2 M solution of hemin by adding 0.456 g to 3.5 ml H$_2$O, to aid solubilisation, add NaOH.
- Add 330 µl of 0.2 M hemin solution to each of the 1 L cultures, resulting in a final concentration of 40 µgml$^{-1}$.
- Incubate overnight at 37 °C.
- Prepare a "metal ion master mix", containing 2 mM Na$^{2+}$, 2 mM Co$^{2+}$, 10 mM Zn$^{2+}$, 10 mM Mn$^{2+}$ and 100 mM Fe$^{3+}$.
- After overnight incubation of cultures, add 1 ml of the metal ion master mix to each of the flasks.
- Incubate at 37 °C for 96 hours (4 days).
- Harvest cells by centrifugation at 6000 rpm for 10 minutes.

**Purification**
- Resuspend cells in 30 mM Tris-HCl pH 8 buffer.
- Lyse cells by 3× passage through French press at 20,000 psi.
- Centrifuge lysate at 11,500 rpm for 45 minutes at 4 °C.
- Retain the supernatant and dialyse overnight against 5 L of ddH$_2$O using 10 kDa molecular weight cut off membrane.
- Following overnight dialysis, add fresh ddH$_2$O every 4 hours for an extra day.

Prepare the following solutions for preparation of a CM-cellulose column and for use during cation exchange chromatography:

1 M NaOH: 8 g NaOH to 200 ml H$_2$O
1.5 M NaCl: 43.83 g to 500 ml H$_2$O
1 M HCl: 17.17 ml to 182.83 ml H$_2$O.
0.1 M NaOH/0.5 M NaCl: 25 ml 1 M NaOH and 83.5 ml 1.5 M NaCl to 141.5 ml H$_2$O.
0.5 M NaOH: 125 ml 1 M NaOH to 125 ml H$_2$O
0.1 M HCl / 0.5 M NaCl: 25 ml 1 M HCl and 83.5 ml 1.5 M NaCl to 141.5 ml H$_2$O
400 mM MES pH 6.0 buffer: 25.59 g to 300 ml H$_2$O. Adjust pH to 6.0 with NaOH.
40 mM MES pH 6.0 buffer: 37.5 ml 400 mM MES buffer to final volume 337.5 ml H$_2$O.
40 mM MES/200 mM NaCl: 12.5 ml 400 mM MES buffer and 16.7 ml 1.5 M NaCl stock to 95.8 ml H$_2$O.

**CM-cellulose column preparation**
- Measure out approximately 125 ml CM-cellulose and add 625 ml ddH$_2$O.
- Allow to settle for 30 – 45 mins, then gently remove the H$_2$O.
- Resuspend in 250 ml of 0.1 M NaOH/0.5 M NaCl, allow to settle for 30 – 45 mins.
- Gently remove 0.1 M NaOH/0.5 NaCl
- Resuspend in 250 ml 0.5 M NaOH, allow to settle for 30 – 45 mins.
- Gently remove 0.5 M NaOH.
- Resuspend in 250 ml 0.1 M HCl/0.5 M NaCl, allow to settle for 30 – 45 mins.
- Gently remove 0.1 M HCl/0.5 M NaCl.
- Add this resin to a gravity-flow chromatography column and wash with 2 column volumes (CV) of H₂O.
- Wash with 2 CVs of 400 mM MES pH 6.0 buffer.
- Wash with 2 CVs of 40 mM MES pH 6.0 buffer.
- Wash with > 5 CVs of ddH₂O.

*Cation exchange chromatography*

- Following 48 hour dialysis of cell extract from AxCytcp overexpression, load onto CM cellulose column.
- Wash with 2 CVs of ddH₂O.
- Wash with 2 CVs of 40 mM MES pH 6.0 buffer. This should result in a brown (wild-type and R124A) or red (L16A) band appearing on the column.
- Prepare a gradient maker containing 125 ml of 40 mM MES pH 6.0 buffer and 125 ml 40 mM MES/200 mM NaCl buffer.
- Collect the elution fractions using a manual fractionator, collecting 5 ml samples.

### 8.5.2 Analysis of fractions

Commonly, purified fractions are analysed by UV-vis spectroscopy, with the ratio between the protein 280 nm band and the Soret, haem-arising 396 nm (wild-type and R124A) or 418 nm (L16A) band. Usually, for wild-type and R124A samples, a ratio of 1:2 in favour of the 396 nm peak is sufficient for spectroscopic studies, whereas for L16A, a ratio of 1:3 in favour of the 418 nm peak is sufficient. Examples of UV-vis spectra of high purity wild-type, L16A and R124A samples are illustrated in Figures 8.7, 8.8 and 8.9, respectively. L16A exhibits a sharp 418 nm peak due to isolation of Cytcp with a CO ligand attached at the 6-coordinate position.
Figure 8.7. UV-vis spectrum of high purity, as isolated wild-type AxCytcp. The Soret peak at 396 is over twice as pronounced as the protein, 280 nm band.

Figure 8.8. UV-vis spectrum of high purity, as isolated L16A AxCytcp. The Soret peak is shifted to 418 nm due to a bound CO ligand, and is over three times as pronounced as the protein, 280 nm band.

Figure 8.9. UV-vis spectrum of high purity, as isolated R124A AxCytcp. The Soret peak at 396 is over twice as pronounced as the protein, 280 nm band.