

INVESTIGATING THE BIOSYNTHESIS OF HAEM  $d_1$   
IN *Pseudomonas aeruginosa*; A COFACTOR FOR  
DISSIMILATORY NITRITE REDUCTASE

---

**Vina Parmar-Bhundia B.Sc (Hons), MRes.**

A thesis submitted to the  
UNIVERSITY OF LONDON  
for the degree of  
DOCTOR OF PHILOSOPHY

## ABSTRACT

---

Haem  $d_1$  is a modified tetrapyrrole unique to the periplasmic enzyme nitrite reductase where it acts in catalysing the reduction of nitrite ( $\text{NO}_2^-$ ) to nitric oxide (NO), as part of denitrification. As with all modified tetrapyrroles, haem  $d_1$  shares a common biosynthetic pathway starting from 5-aminolaevulinic acid (ALA), up to the formation of uroporphyrinogen III (UIII). UIII is the branch point from which the pathway diverges to form the various metallo-prosthetic groups including vitamin B<sub>12</sub>.

The precise mechanism of transformation from UIII to haem  $d_1$  is unknown. Examination of both structures shows a requirement of methylation at C2 and C7; decarboxylation of acetate side chains at C12 and C18; loss of propionic side chains at C3 and C8 with subsequent oxidation at C3 and C8; dehydrogenation of C17 propionate side chain gives the acrylate substituent and ferrochelation. Of particular interest is the addition of oxygen to the macrocycle under anaerobic conditions. Only one other intermediate, compound 800, has been isolated thus far but it is unknown how it is part of the pathway. Genetic studies have implicated seven *nir* genes, called *nirF*, *nirD*, *nirL*, *nirG*, *nirH*, *nirJ* and *nirE*, are required for haem  $d_1$  biogenesis.

Here, experiments and data show for the first time that it proceeds from UIII to precorrin-2 using the enzyme NirE. This study is the first to experimentally show the production of precorrin-2 as part of the pathway using anaerobic enzyme assays.

This thesis illustrates the intense work that has focused on cloning the genes individually and as multigene constructs in an attempt to characterise the proteins overproduced. Heterologous expression in *Escherichia coli* has been successful as well as the development of a homologous expression system in *Pseudomonas aeruginosa*. The data represented shows the various aspects entailed in the optimisation of overproduction and the stabilisation of the Nir proteins. It also documents the first concerted attempt to take the operon and engineer strains to make haem  $d_1$  both *in vivo* and *in vitro*, using the Link and Lock method to clone the *nir* genes consecutively into a plasmid. This thesis therefore provides a foundation for understanding the molecular biology and biochemistry of haem  $d_1$  synthesis for the future.

# CONTENTS

---

Abstract.....	i
Contents.....	ii
List of Figures.....	viii
List of Tables.....	xii
Abbreviations.....	xiii
Acknowledgments.....	xvi

## CHAPTER 1: AN INTRODUCTION TO TETRAPYRROLES, THEIR SYNTHESIS IN *PSEUDOMONADS* AND THE ENIGMA THAT IS HAEM $d_1$ .

1. Introduction.....	2
1.1. Modified tetrapyrroles.....	2
1.1.1. Chlorophyll and Bacteriochlorophyll.....	5
1.1.2. Haem.....	7
1.1.3. Sirohaem.....	9
1.1.4. Coenzyme F <sub>430</sub> .....	12
1.1.5. Vitamin B <sub>12</sub> .....	13
1.2. Haem $d_1$ .....	15
1.2.1. Haem $d_1$ in cytochrome $cd_1$ .....	17
1.2.2. Denitrification.....	20
1.2.3. The denitrification enzyme system.....	21
1.3. The biosynthesis of Haem $d_1$ .....	23
1.3.1. Synthesis of 5-aminolaevulinic acid (ALA).....	23
1.3.2. Synthesis of the first macrocyclic intermediate, uroporphyrinogen III (U <sub>III</sub> ) from 5-aminolaevulinic acid (ALA).....	25
1.3.3. The methylation of U <sub>III</sub> and the formation of PC-2.....	27
1.3.3.1. CobA in the vitamin B <sub>12</sub> pathway.....	28
1.3.3.2. CysG <sup>A</sup> in <i>E. coli</i> and <i>S. enterica</i> in the sirohaem pathway.....	29
1.3.3.3. SirA in <i>Bacilli</i> in the sirohaem pathway.....	31
1.3.3.4. Met1p in <i>Saccharomyces cerevisiae</i> in the sirohaem pathway.....	31

pathway.....	31
1.3.3.5. CorA in <i>Methanobacterium ivanovii</i> .....	31
1.3.3.6. UPM1 from <i>A. thaliana</i> and ZmSUMT1 from maize .....	32
1.3.4. The biosynthesis of haem $d_1$ .....	33
1.3.4.1. NirE as a SUMT.....	35
1.3.4.2. The enigma of haem $d_1$ biosynthesis .....	37
1.4. <i>Ps. aeruginosa</i> as a candidate species .....	39
1.5. Purpose of this thesis.....	40
<b>2. CHAPTER 2: MATERIALS AND METHODS.</b>	
2.1. Materials.....	42
2.1.1. Chemicals (source).....	42
2.1.2. Bacterial Strains .....	42
2.1.3. Plasmids.....	42
2.1.4. Media and Solutions for bacterial work .....	54
2.1.5. Solutions for DNA work.....	57
2.1.6. Solutions for Protein work.....	58
2.1.6.1. Solutions for metal chelate chromatography.....	58
2.1.6.2. Solutions for SDS-PAGE and Native gels .....	60
2.1.6.3. Solutions for Western Blotting.....	62
2.1.6.4. Solutions for Gel Filtration Chromotography (FPLC).....	64
2.2. Molecular Biology Methods.....	65
2.2.1. Isolation of genomic DNA from <i>Ps. aeruginosa</i> .....	65
2.2.2. Isolation of plasmid DNA.....	65
2.2.3. Restriction of DNA .....	66
2.2.4. Electrophoresis of DNA.....	66
2.2.4.1. Agarose gel.....	66
2.2.4.2. Visualisation and UV .....	66
2.2.4.3. Hyperladder (Bioline).....	67
2.2.5. Isolation and purification of a DNA fragment from an Agarose gel.....	67
2.2.6. Ligation of DNA fragments .....	67
2.2.7. Competent cells and Transformation.....	68

2.2.7.1.	Preparation of <i>E. coli</i> Competent Cells.....	68
2.2.7.2.	Transformation of <i>E. coli</i> Competent Cells.....	68
2.2.7.3.	Preparation of <i>Ps. aeruginosa</i> Competent cells.....	68
2.2.7.4.	Transformation of <i>Ps. aeruginosa</i> Competent cells.....	69
2.2.7.5.	Preparation of <i>S. enterica</i> AR3612 Competent cells.....	69
2.2.7.6.	Transformation of <i>S. enterica</i> AR3612 Competent cells.....	70
2.2.7.7.	<i>S. enterica cysG</i> complementation.....	70
2.2.7.8.	Glycerol stocks of strains.....	70
2.2.8.	Polymerase Chain Reaction.....	71
2.2.8.1.	PCR of <i>nir</i> genes.....	71
2.2.8.1.1.	<i>NdeI</i> and <i>BamHI</i> Primers.....	71
2.2.8.1.2.	<i>NdeI</i> or <i>NheI</i> , RBS and <i>XbaI</i> and <i>BamHI</i> Primers.....	72
2.2.8.2.	Primers for PCR.....	73
2.2.9.	T-Vector cloning of PCR products.....	74
2.2.10.	Subcloning into pET14b: <i>NdeI</i> – <i>BamHI</i> for Protein Expression in pET14b.....	75
2.2.11.	Subcloning into pUCP <i>NdeI/Nco</i> for expression in <i>Ps. aeruginosa</i> .....	76
2.2.12.	Link and Lock construction of operon.....	78
2.2.13.	Restriction analysis of <i>nir</i> genes in haem <i>d</i> <sub>1</sub> operon.....	81
2.3.	Biochemical Methods.....	82
2.3.1.	Protein assay (Bradford Assay).....	82
2.3.2.	A <sub>280</sub> protein concentration estimation.....	82
2.3.3.	Standard cloning, growth and purification of histagged proteins.....	83
2.3.3.1.	Protein overproduction in <i>E.coli</i> .....	83
2.3.3.2.	Protein overproduction in <i>Ps. aeruginosa</i> .....	83
2.3.3.3.	Sonication of bacteria.....	84
2.3.3.4.	Purification by Metal chelate chromatography.....	84
2.3.3.5.	Histag cleavage.....	85
2.3.3.6.	Buffer Exchange / Desalting.....	85
2.3.3.7.	Gel filtration / Fast protein liquid chromatography.....	85
2.3.4.	Poly-Acrylamide Gel Electrophoresis (PAGE) of proteins.....	86
2.3.4.1.	SDS-PAGE of proteins.....	86
2.3.4.2.	Native-PAGE of proteins.....	86

2.3.5. Western blotting of histagged proteins.....	87
2.3.5.1. SDS-PAGE for Western blotting of proteins.....	87
2.3.5.2. Western blotting.....	87
2.3.5.3. Probing for Histagged proteins.....	88
2.4. Isolation and production of tetrapyrroles.....	89
2.4.1. UltraViolet-Visible Spectroscopy.....	89
2.4.2. Production of precorrin-2 and sirohydrochlorin <i>in situ</i> from cell extract.....	89
2.4.3. Production of Precorrin-2 and sirohydrochlorin from purified protein.....	90
2.4.4. Production of haem $d_1$ intermediates.....	91
2.4.4.1. <i>In vivo</i> accumulation of intermediates.....	91
2.4.4.2. <i>In vitro</i> accumulation of intermediates.....	91
2.4.5. <i>In vitro</i> Sirohydrochlorin coupled SUMT assay.....	92
2.4.6. Reverse Phase chromatography Separation of Pigments.....	93
2.4.7. Ion exchange chromatography separation of pigments – DEAE Sephacyl.....	93
2.5. Crystallography - Hanging drop method.....	94
3. CHAPTER 3: CHARACTERISATION OF NirE.	
3.1. Introduction.....	96
RESULTS AND DISCUSSION.....	102
3.2. Sequence identity of NirE to various SUMTs.....	102
3.3. The overproduction of NirE.....	107
3.3.1. Cloning and expression of <i>nirE</i> in <i>E.coli</i> .....	107
3.3.2. Cloning and expression of <i>nirE</i> in <i>Ps. aeruginosa</i> .....	107
3.3.3. Purification of recombinant NirE from <i>E. coli</i> and <i>Ps. aeruginosa</i> .....	108
3.3.4. Analysis of recombinant NirE from <i>E. coli</i> and <i>Ps. aeruginosa</i>	
3.3.5. The accumulation of porphyrinoid material.....	108

3.4. Complementation of <i>Salmonella enterica cysG metH</i> mutant strains by <i>Ps. aeruginosa</i> NirE.....	115
3.5. Biochemical complementation of NirE.....	118
3.5.1. Production of PC-2 and SHC <i>in situ</i> using NirE.....	118
3.5.2. SHC coupled NirE assay.....	123
3.6. Conclusion.....	129
4. CHAPTER 4: OPTIMISATION OF PRODUCTION AND STABILISATION OF PROTEINS INVOLVED IN HAEM $d_1$ SYNTHESIS.	
4.1. Introduction.....	133
RESULTS AND DISCUSSION.....	136
4.2. Approaches employed in the analysis of NirF, D, L, G, H, J.....	136
4.2.1. Analysis of rare codons within the <i>nir genes</i> .....	136
4.2.2. Cloning and heterologous overproduction of proteins in <i>E. coli</i> .....	137
4.2.2.1. Protease inhibitors.....	137
4.2.3. Control of expression by control of [IPTG] within a host strain.....	138
4.2.4. Cloning & homologous overproduction of proteins in <i>Ps. aeruginosa</i> .....	138
4.2.5. Protein stabilisation screen.....	139
4.3. NirD, NirL, NirG, NirH.....	141
4.3.1. Bioinformatic analysis.....	141
4.3.2. Optimisation of overproduction.....	144
4.3.2.1. Overproduction in <i>E. coli</i> .....	144
4.3.2.2. Overproduction in <i>Ps. aeruginosa</i> .....	150
4.3.2.3. Co-expression of NirD, NirL, NirG, and NirH.....	152
4.3.3. Stabilisation screen.....	155
4.3.4. NirG crystallisation trials.....	157
4.4. NirF.....	158
4.5. NirJ.....	162
4.6. Conclusion.....	166

<b>5. CHAPTER 5: A SYNTHETIC BIOLOGY APPROACH TOWARDS THE GENERATION OF TETRAPYRROLE INTERMEDIATES OF THE HAEM <math>d_1</math> PATHWAY <i>IN VIVO</i> AND <i>IN VITRO</i>.</b>	
5.1. Introduction.....	170
5.2. <i>In vivo</i> accumulation of intermediates.....	173
5.2.1. The expression of the operon in <i>E.coli</i> .....	173
5.2.2. Link and Lock cloning of nir genes into an operon.....	174
5.2.3. Production, extraction and separation of accumulated tetrapyrroles.....	177
5.2.4. Results obtained from expression in <i>E. coli</i> JM109.....	178
5.2.5. Results obtained from expression in <i>Ps. aeruginosa</i> .....	180
5.2.6. Discussion of data obtained from expression of the haem $d_1$ operon in <i>E. coli</i> and <i>Ps. aeruginosa</i> .....	182
5.3. <i>In vitro</i> expression of Intermediates.....	184
5.3.1. Production, extraction and separation of accumulated tetrapyrroles.....	184
5.3.2. Results obtained from incubation of Nir proteins with PC-2 and SHC produced <i>in situ</i> .....	187
5.3.3. Results obtained from incubation of Nir proteins with PC-2 and SHC produced <i>in vitro</i> .....	192
5.3.4. Discussion of data obtained from incubation of cell extracts containing overproduced Nir proteins with PC-2 and SHC produced <i>in situ</i> or <i>in vitro</i> .....	195
5.4. Conclusion.....	197
<b>6. CHAPTER 6: CONCLUSIONS.....</b>	<b>201</b>

## LIST OF FIGURES

---

Figure 1.1: The structure of Uroporphyrinogen III (UIII).....	3
Figure 1.2: Summary of Tetrapyrrole biosynthesis: a branched pathway from Uroporphyrinogen III (UIII).....	3
Figure 1.3: The structures of Chlorophyll <i>a</i> and Bacteriochlorophyll <i>a</i> .....	6
Figure 1.4: The structure of Haem <i>b</i> .....	6
Figure 1.5: The haem <i>b</i> biosynthetic pathway.....	8
Figure 1.6: The structure of Sirohaem.....	10
Figure 1.7: The role of Sirohaem in cysteine biosynthesis .....	10
Figure 1.8: Diversity in the biosynthetic pathway of the transformation of Uroporphyrinogen III into Sirohaem in different organisms.....	11
Figure 1.9: The structure of Coenzyme F <sub>430</sub> .....	12
Figure 1.10: The structure of Vitamin B <sub>12</sub> .....	14
Figure 1.11: The structure of Haem <i>d</i> <sub>1</sub> .....	16
Figure 1.12 Cytochrome <i>cd</i> <sub>1</sub> structure .....	18
Figure 1.13: $\beta$ -propeller <i>d</i> <sub>1</sub> domain structure .....	19
Figure 1.14: Topology of the denitrification system in all denitrifying bacteria. ....	22
Figure 1.15: Biosynthesis of ALA via the C-4 Shemin route.....	24
Figure 1.16: Biosynthesis of ALA via the C-5 route .....	24
Figure 1.17: The biosynthetic conversion of ALA to UIII.....	26
Figure 1.18: Summary of the tetrapyrrole pathway from UIII.....	26
Figure 1.19: The transformation of UIII to PC-2 via PC-1.....	27
Figure 1.20: The structure of CobA from <i>Ps. denitrificans</i> .....	28
Figure 1.21: The structure of CysG from <i>S. enterica</i> .....	30
Figure 1.22: The organisation of haem <i>d</i> <sub>1</sub> genes within the operon. ....	34
Figure 1.23: The transformation from UIII to form haem <i>d</i> <sub>1</sub> .....	36
Figure 2.1: The promoter and multiple cloning sequences of pGEMT and pGEMT-Easy Vectors.....	75
Figure 2.2: The promoter and multiple cloning sites of pET14b.....	76

<b>Figure 2.3: The promoter and multiple cloning sites of pUCP-Nde / pUCP-Nco .....</b>	<b>77</b>
<b>Figure 2.4: <i>XbaI</i> and <i>NheI</i>; the formation of compatible cohesive ends .....</b>	<b>79</b>
<b>Figure 2.5: Diagrammatic representation of link and lock method of cloning .....</b>	<b>80</b>
<b>Figure 2.6: Assembly of the western blotting apparatus .....</b>	<b>88</b>
<b>Figure 3.1: SUMTs and their associated downstream enzymes – a diagrammatic representation.....</b>	<b>97</b>
<b>Figure 3.2: Alignment of NirE, <i>Ps. aeruginosa</i> with CobA. <i>Ps. denitrificans</i> .....</b>	<b>103</b>
<b>Figure 3.3: Multiple sequence alignment of known characterised SUMT enzymes .....</b>	<b>104</b>
<b>Figure 3.4: The structure of pyocyanin.....</b>	<b>109</b>
<b>Figure 3.5: SDS-PAGE of NirE from <i>E.coli</i>. .....</b>	<b>110</b>
<b>Figure 3.6: SDS-PAGE of NirE from <i>Ps. aeruginosa</i>.....</b>	<b>110</b>
<b>Figure 3.7: SDS-PAGE of NirE from <i>Ps. aeruginosa</i>.....</b>	<b>111</b>
<b>Figure 3.8: UV-VIS Spectra of NirE with porphyrinoid material.....</b>	<b>111</b>
<b>Figure 3.9: The structure of trimethylpyrrocorphin .....</b>	<b>113</b>
<b>Figure 3.10: NirE cells containing trimethylpyrrocorphin fluoresce red under UV light. ....</b>	<b>113</b>
<b>Figure 3.11: The characteristic red pigment of SUMTs shown by NirE. ....</b>	<b>114</b>
<b>Figure 3.12: Spectra showing the transformation of PC-2 to SHC by CobA. ....</b>	<b>119</b>
<b>Figure 3.13: <i>In situ</i> production of PC-2 and SHC with CobA. ....</b>	<b>120</b>
<b>Figure 3.14: <i>In situ</i> production of PC-2 with NirE. ....</b>	<b>121</b>
<b>Figure 3.15: <i>In situ</i> production of SHC with NirE.....</b>	<b>122</b>
<b>Figure 3.16: UV-Vis Spectra showing the <i>in vitro</i> generation of SHC by NirE .....</b>	<b>125</b>
<b>Figure 3.17: The rate of generation of SHC by NirE.....</b>	<b>126</b>
<b>Figure 3.18: UV-Vis Spectra of reaction mixture without SAM .....</b>	<b>127</b>
<b>Figure 3.19: UV-Vis Spectra of reaction mixture without NAD<sup>+</sup> .....</b>	<b>128</b>
<b>Figure 3.20: The first step of haem <i>d</i><sub>1</sub> biosynthesis.....</b>	<b>131</b>
<b>Figure 4.1: The haem <i>d</i><sub>1</sub> operon.....</b>	<b>135</b>
<b>Figure 4.2: Sequence alignment of NirD, NirL, NirG, NirH to LRP from <i>E. coli</i>.....</b>	<b>142</b>

<b>Figure 4.3: Sequence alignments of NirD, NirL, NirG, NirH. ....</b>	<b>143</b>
<b>Figure 4.4: SDS-PAGE gels of incubation at 37 °C and 16 °C for NirL. ....</b>	<b>145</b>
<b>Figure 4.5: Gel filtration elution profiles of NirL, NirG and NirH. ....</b>	<b>146</b>
<b>Figure 4.6: SDS-PAGE of NirL at 0.1mM IPTG 0.04mM IPTG in Tuner Cells .....</b>	<b>147</b>
<b>Figure 4.7 SDS-PAGE gel of NirD incubated at 37 °C and 16 °C. ....</b>	<b>148</b>
<b>Figure 4.8: NirD gel filtration elution profile and SDS-PAGE of eluted fractions. ....</b>	<b>149</b>
<b>Figure 4.9: NirG, NirH and NirD western blots from overproduction in <i>Ps. aeruginosa</i> aerobically and anaerobically. ....</b>	<b>151</b>
<b>Figure 4.10: SDS-PAGE analysis of the co production of NirD, NirL, NirG and NirH in <i>E. coli</i> (BL21 DE3 pLysS) . ....</b>	<b>153</b>
<b>Figure 4.11: SDS-PAGE of NirG in HEPES. ....</b>	<b>155</b>
<b>Figure 4.12: Elution profile of NirG in HEPES . ....</b>	<b>156</b>
<b>Figure 4.13: Non-denaturing (native) gels of NirG of the corresponding fractions from gel filtration chromatography . ....</b>	<b>156</b>
<b>Figure 4.14: Crystal formations in 60 % MPD, 0.1 M Tris pH 8.5 and 0.2 M ammonium phosphate. ....</b>	<b>157</b>
<b>Figure 4.15: Comparison of sequences of NirF and NirE . ....</b>	<b>159</b>
<b>Figure 4.16: SDS-PAGE analysis of NirF with induction at 16 °C and 37 °C and in the presence of protease inhibitors . ....</b>	<b>160</b>
<b>Figure 4.17: Elution profile of NirF with gel filtration chromatography . ....</b>	<b>160</b>
<b>Figure 4.18: NirF western blots from overproduction in <i>Ps. aeruginosa</i> aerobically and anaerobically . ....</b>	<b>161</b>
<b>Figure 4.19: Multiple sequence alignment of NirJ. ....</b>	<b>163</b>
<b>Figure 4.20: SDS-PAGE and western blot of NirJ in <i>E. coli</i>, BL21 DE3 pLysS and <i>Ps. aeruginosa</i> (aerobic and anaerobic) . ....</b>	<b>165</b>
<b>Figure 4.21: The steps of haem <math>d_1</math> biosynthesis . ....</b>	<b>168</b>
<b>Figure 5.1: The pathway from UIII to haem <math>d_1</math> . ....</b>	<b>171</b>
<b>Figure 5.2: SDS-PAGE of the haem <math>d_1</math> operon overexpressed in <i>E. coli</i>. ....</b>	<b>174</b>
<b>Figure 5.3: Link and Lock method of cloning . ....</b>	<b>176</b>
<b>Figure 5.4: SDS – PAGE analysis of the expression of the operon constructed with ribosome binding sites (RBS) . ....</b>	<b>177</b>
<b>Figure 5.5: UV-Vis spectra of all isolated supernatants from <i>E. coli</i> prior to separation on RP-18 columns. ....</b>	<b>178</b>

<b>Figure 5.6: UV-Vis spectra of all isolated fractions from <i>E. coli</i> after separation on RP-18 columns.....</b>	<b>179</b>
<b>Figure 5.7: UV-Vis spectra of all aerobically isolated supernatants from <i>Ps. aeruginosa</i> prior to separation on RP-18 columns.....</b>	<b>180</b>
<b>Figure 5.8: UV-Vis spectra of all aerobically isolated supernatants from <i>Ps. aeruginosa</i> after separation on RP-18 columns.....</b>	<b>180</b>
<b>Figure 5.9: UV-Vis spectra of all anaerobically isolated supernatants from <i>Ps. aeruginosa</i> prior to separation on RP-18 columns.....</b>	<b>181</b>
<b>Figure 5.10: UV-Vis spectra of all anaerobically isolated supernatants from <i>Ps. aeruginosa</i> after separation on RP-18 columns.....</b>	<b>181</b>
<b>Figure 5.11: UV-Vis spectra of NirF, D, L, G, H, J, incubated anaerobically with PC-2 and SHC with all cofactors. ....</b>	<b>187</b>
<b>Figure 5.12: UV-Vis spectra of NirF, D, L, G, H, J incubated anaerobically with PC-2 and SHC, cell extracts of <i>Ps. aeruginosa</i> and <i>Pa. pantrophus</i> and all cofactors. ....</b>	<b>188</b>
<b>Figure 5.13: UV-Vis spectra of cell extracts of <i>Ps. aeruginosa</i>, <i>E. coli</i> and <i>Pa. pantrophus</i> incubated anaerobically with PC-2 and SHC and all cofactors (No Nir proteins).....</b>	<b>189</b>
<b>Figure 5.14: UV-Vis spectra of anaerobic incubations with PC-2 using cell extracts of each overproduced Nir protein separately.....</b>	<b>190</b>
<b>Figure 5.15: UV-Vis spectra of anaerobic incubations with SHC using cell extracts of each overproduced Nir protein separately.....</b>	<b>191</b>
<b>Figure 5.16: UV-Vis spectra PC-2 and SHC produced from purified HemB, HemC, HemD, CobA <math>\pm</math> SirC respectively (<math>\pm</math>NAD<sup>+</sup>).....</b>	<b>192</b>
<b>Figure 5.17: UV-Vis spectra of compounds eluted at 1M NaCl from anaerobic incubation of PC-2 with NirFDLGHJ with varying cofactors (39-41).....</b>	<b>193</b>
<b>Figure 5.18: UV-Vis spectra of compounds eluted at 1M NaCl from the anaerobic incubation of SHC with NirFDLGHJ and all cofactors (42-44)..</b>	<b>194</b>
<b>Figure 6.1: Schematic representation of the putative pathway to haem <i>d</i><sub>1</sub> ..</b>	<b>207</b>

## LIST OF TABLES

---

<b>Table 2.1: List of stock strains used in this study .....</b>	<b>43</b>
<b>Table 2.2: List of strains used as part of this study .....</b>	<b>44</b>
<b>Table 2.3: List of stock vectors used in the study.....</b>	<b>47</b>
<b>Table 2.4: List of plasmids constructed and used in this study .....</b>	<b>48</b>
<b>Table 2.5: Composition of the running gels made to the desired concentration of acrylamide depending on the size of the protein.....</b>	<b>60</b>
<b>Table 2.6: Composition of the stacking gels made to the desired concentration .....</b>	<b>60</b>
<b>Table 2.7: Amino acids listed with their properties .....</b>	<b>65</b>
<b>Table 2.8: Tabulation of all primers used in this study to clone the genes individually and for the Link and Lock method. ....</b>	<b>73</b>
<b>Table 2.9: Restriction enzymes unique to each <i>nir</i> gene and plasmid.....</b>	<b>81</b>
<b>Table 3.1: A summary of SUMT amino acid homology .....</b>	<b>106</b>
<b>Table 3.2: Summary of results obtained from complementation of <i>S. enterica cysG metE</i> (AR3612).....</b>	<b>117</b>
<b>Table 4.1: Tabulation of rare codon usage for each <i>nir</i> gene.....</b>	<b>136</b>
<b>Table 4.2: Comparison of the amino acid sequences from NirF, NirD, NirL, NirG, NirH and NirJ sequences of <i>Ps. aeruginosa</i> (PAO) with <i>Ps. stutzeri</i> and <i>Pa. denitrificans</i>, using Clustal W.....</b>	<b>142</b>
<b>Table 4.3: Comparison of soluble protein yields from NirD, NirL, NirG and NirH from 1L cultures incubated at 37 °C and 16 °C. ....</b>	<b>145</b>
<b>Table 5.1: Plasmids constructed for the accumulation of haem <math>d_1</math> intermediates .....</b>	<b>175</b>
<b>Table 5.2: Summary of incubations of Nir proteins with PC-2 and SHC.....</b>	<b>185</b>
<b>Table 6.1: Summary of possible roles of the Nir Proteins.....</b>	<b>208</b>

## ABBREVIATIONS

---

Å	Angstrom
A <sub>280</sub>	Absorbance at 280 nm
Abs	Absorbance
A	Adenine
ADP	Adenosine diphosphate
ALA	5-Aminolaevulinic Acid
Amp	Ampicillin
Amp <sup>R</sup>	Ampicillin resistance
ANR	Anaerobic regulator of arginine deaminase and nitrate reductase
APS	Ammonium persulfate
ATP	Adenosine triphosphate
B <sub>12</sub>	Vitamin B <sub>12</sub> , Cobalamin
C	cytosine
Cb	Carbenicillin
Cb <sup>R</sup>	Carbenicillin resistance
Cm	Chloramphenicol
Cm <sup>R</sup>	Chloramphenicol resistance
Da	Dalton
λDE3	Bacteriophage lambda carrying the gene for T7 RNA polymerase
DEAE	Diethylaminoethyl sepharose
dH <sub>2</sub> O(st)	Distilled water (sterile)
DNA	Deoxyribonucleic acid
DNR	Dissimilatory nitrite reductase
dNTP	2'-Deoxy-nucleoside (A, C, G, T)-5'-triphosphate
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiaminetetraacetic acid
e.g.	For example
F'	Host contains an F' episome, with the stated features
FNR	Anaerobic regulator of fumarate and nitrate reductase
FPLC	Fast Protein Liquid Chromatography

G	Guanine
HMB	Hydroxymethylbilane
HPLC	High Performance Liquid Chromatography
HPLC-MS	High Performance Liquid Chromatography coupled to Mass Spectroscopy
HemB	ALA dehydratase
HemC	Porphobilinogen deaminase
HemD	Uroporphyrinogen III synthase
<i>hemB</i>	ALA dehydratase gene
<i>hemC</i>	Porphobilinogen deaminase gene
<i>hemD</i>	Uroporphyrinogen III synthase gene
HEPES	4-(2-Hydroxyethyl) piperazine-1-ethanesulfonic acid
IPTG	Isopropyl- $\beta$ -D-thiogalactoside
i.e.	Id est (That Is)
kDa	Kilo Dalton
<i>lac</i>	Lactose promoter
LB	Luria-Bertani medium
LC	Liquid chromatography
pLysS	LysS plasmid
M	Molar
MCS	Multiple cloning site
MS	Mass spectometry
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide hydrate
NADPH	Nicotinamide adenine dinucleotide 3'-phosphate
NaR	Nitrate reductase
NEB	New England Biolabs
NiR	Nitrite reductase
NNR	Anaerobic regulator of nitric oxide and nitrate reductase
nm	Nanometer
NoS	Nitric oxide synthase
NoR	Nitric oxide reductase
NO <sub>3</sub> <sup>-</sup>	Nitrate
NO <sub>2</sub> <sup>-</sup>	Nitrite
NO*	Nitric oxide

N <sub>2</sub> O	Nitrous oxide
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PAO1	<i>wt Pseudomonas aeruginosa</i>
<i>Pa. denitrificans</i>	<i>Paracoccus denitrificans</i>
<i>Pa. pantotrophus</i>	<i>Paracoccus pantotrophus</i>
<i>Ps. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
<i>Ps. stutzeri</i>	<i>Pseudomonas stutzeri</i>
<i>Ps. denitrificans</i>	<i>Pseudomonas denitrificans</i>
PBG	Porphobilinogen
PCR	Polymerase chain reaction
RBS	Ribosome binding site
RNA	Ribonucleic acid
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine
SDS	Sodium dodecyl sulphate
SP6	Promoter
<i>S. enterica</i>	<i>Salmonella enterica subsp. enterica serovar typhimurium strain T</i>
SUMT	S-Adenosyl-L-methionine uroporphyrinogen III methyltransferase
T	Thymine
T7P	T7 Promoter
T7T	T7 Terminator
tac	Tryptophan and lactose hybrid promoter
TAE	Tris-acetate-EDTA Buffer
TCA	Trichloroacetic Acid
TE	Tris-EDTA Buffer
TEMED	N,N,N,N -Tetramethyl-ethylenediamine
Tris	Tris-(hydroxymethyl)-aminomethane base
UIII	Uroporphyrinogen III
UV-Vis	Ultraviolet - visible
v/v	Volume by volume
w/v	Weight by volume
wt	Wild type
X-gal	5-Bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

## ACKNOWLEDGEMENTS

---

First and foremost I would like to thank Professor Martin Warren, my supervisor, for his constant patience, support, faith and kind words; all of which I needed and had a lack of! You have been so kind and generous I thank you from the bottom of my heart for everything. I am just so sorry I could not finish sooner but here I am!

My mum, who is no longer with us, I miss you dearly even now. I wish you were here but I am glad you are not in any pain anymore. I am sorry I could not do more for you but I have finished my studies just like I promised you and I hope you are looking down with a huge smile and blessing all your children and grandchildren and my father. Thank you seems so simple a word for all you have done for us with Dad. Dad, thank you for your sacrifices, help, generosity and encouragement, you have made us stronger by being our rock after mum passed way. I hope I have made you both proud.

Thank you to my dear husband, Ajay, who probably feels he has written this thesis too! Thanks for being the person that you are and though I may not seem like I do often, I do truly appreciate all that you have sacrificed for me. I know it has not been easy in any way but it can only get better and better, just the two of us. I look forward to a thesis free future!!

My sister, Mina, and my brother, Rakesh, I thank you for being there for our mum. I wish I could have more than I was able to but it was not to be. Thanks for understanding and being there for me too. I will never forget it. Of course I can't forget my niece and nephew, Aanya and Ashish, you gave me a reason to smile and forget my woes when we went to the park and played. You truly showed me how to smile again and reminded me of a saying:

***“Seek the wisdom of the ages, but look at the world through the eyes of a child” (R.Wild)***

I would like to thank God for his kind and those not so kind blessings that make me the person am I today.

Thanks to all those in the lab, Dr Evelyne Deery, Dr Amanda Brindley (Mo), Dr Helen Leech, Dr Dana Heldt, Dr Stephanie Frank (Steffi), Dr Andrew Lawrence, Dr Ruth-Sarah Rose (Ru), and those who I may have crossed paths with along the way. I have some wonderful memories of my times in the lab and miss them too! Evelyne, thanks for all the help with the cloning and for the constant support you gave. Mo thanks for your patience with all the protein work and nice that we finally found some terms to agree on in the end, albeit the last week!! Helen, thanks again for your support and help with the glove box work as it was appreciated. Andrew, Ru, Dana and Steffi, you always made me laugh and made my time in the lab memorable. Dana and Steffi, thanks for being good friends when I needed them, and thanks for being a constant support. I want to mention Rebecca; it was lovely to get to know you even for a short time and those starbucks coffees at the weekend! I hope everything everyone has wished for works out for them.

***I dedicate this thesis to the memory of my mother, Sushila Parmar, an inspirational mum, wife and grandmother.***

**An Introduction to Tetrapyrroles,  
their Synthesis in *Pseudomonads*  
and the Enigma that is Haem  $d_1$**

## 1. Introduction

### 1.1. Modified tetrapyrroles

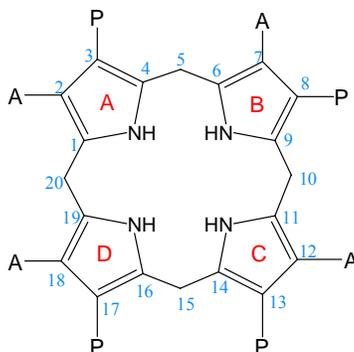
Molecules such as chlorophyll, haem, vitamin B<sub>12</sub>, factor F<sub>430</sub>, sirohaem and haem  $d_1$  pertain to a family of compounds termed modified tetrapyrroles. They are utilised in nature as the prosthetic groups of different enzymes that are part of fundamental processes. Each modified tetrapyrrole fulfils important diverse chemical and biological functions facilitated by the central metal ion moiety and its peripheral groups. These properties also give them distinct colours such as the red of haemoglobin and green in chlorophyll and are thus referred to as being “the pigments of life”.

As the name suggests tetrapyrroles are macrocyclic structures composed for four pyrrole units, each being a pentagonal ring structure to form the basic porphyrin framework. All tetrapyrroles share a common biosynthetic pathway up to the formation of uroporphyrinogen III (UIII), the first macrocyclic intermediate, as depicted in Figure 1.1. From this point onwards the pathway diverges to form the various tetrapyrrole derivatives through unique modification of peripheral groups and chelation of a different metal ion at its core.

For brevity, the synthesis of UIII will be discussed in detail later. Starting from 5 - aminolaevulinic acid (ALA), two molecules of ALA condense to form porphobilinogen (PBG), catalysed by ALA dehydratase. Four PBG molecules then condense to form a linear tetrapyrrole, hydroxymethylbilane (HMB), also called pre-uroporphyrinogen, in the presence of PBG deaminase. This compound then cyclises to form UIII, the precursor to various modified tetrapyrroles, as summarised in Figure 1.2.

**Figure 1.1: The structure of Uroporphyrinogen III (UIII).**

UIII is the first macrocyclic intermediate and a key branch point molecule in the pathway to various tetrapyrroles. Decarboxylation of its acetate groups directs flux toward haem and chlorophyll synthesis whereas methylation of the macrocycle initiates synthesis toward sirohaem, vitamin B<sub>12</sub>, and haem  $d_1$ . The systematic numbering of the tetrapyrrole carbons is shown in blue and the pyrrole groups are labelled, alphabetically, in red. A = CH<sub>2</sub>CO<sub>2</sub>H, P = CH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H.

**Figure 1.2: Summary of Tetrapyrrole biosynthesis: a branched pathway from Uroporphyrinogen III (UIII).**

This illustrates the shared pathway from ALA to UIII of all modified tetrapyrroles. UIII acts as a precursor to form haem, chlorophyll, vitamin B<sub>12</sub>, sirohaem, coenzyme F<sub>430</sub> and haem  $d_1$ . UIII via protoporphyrin IX yields haem and chlorophyll. The formation of sirohaem requires the conversion of UIII to precorrin 2 (PC2), then sirohydrochlorin (SHC) and chelation. Vitamin B<sub>12</sub> is formed from PC-2 directly. The transformation of PC-2 to yield factor F<sub>430</sub> and Haem  $d_1$  remain as yet to be elucidated. For brevity, the pathway to each of these has not been shown in detail but one must appreciate the complicated steps required in the transformation of UIII to each modified tetrapyrrole. A = CH<sub>2</sub>CO<sub>2</sub>H, P = CH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H.

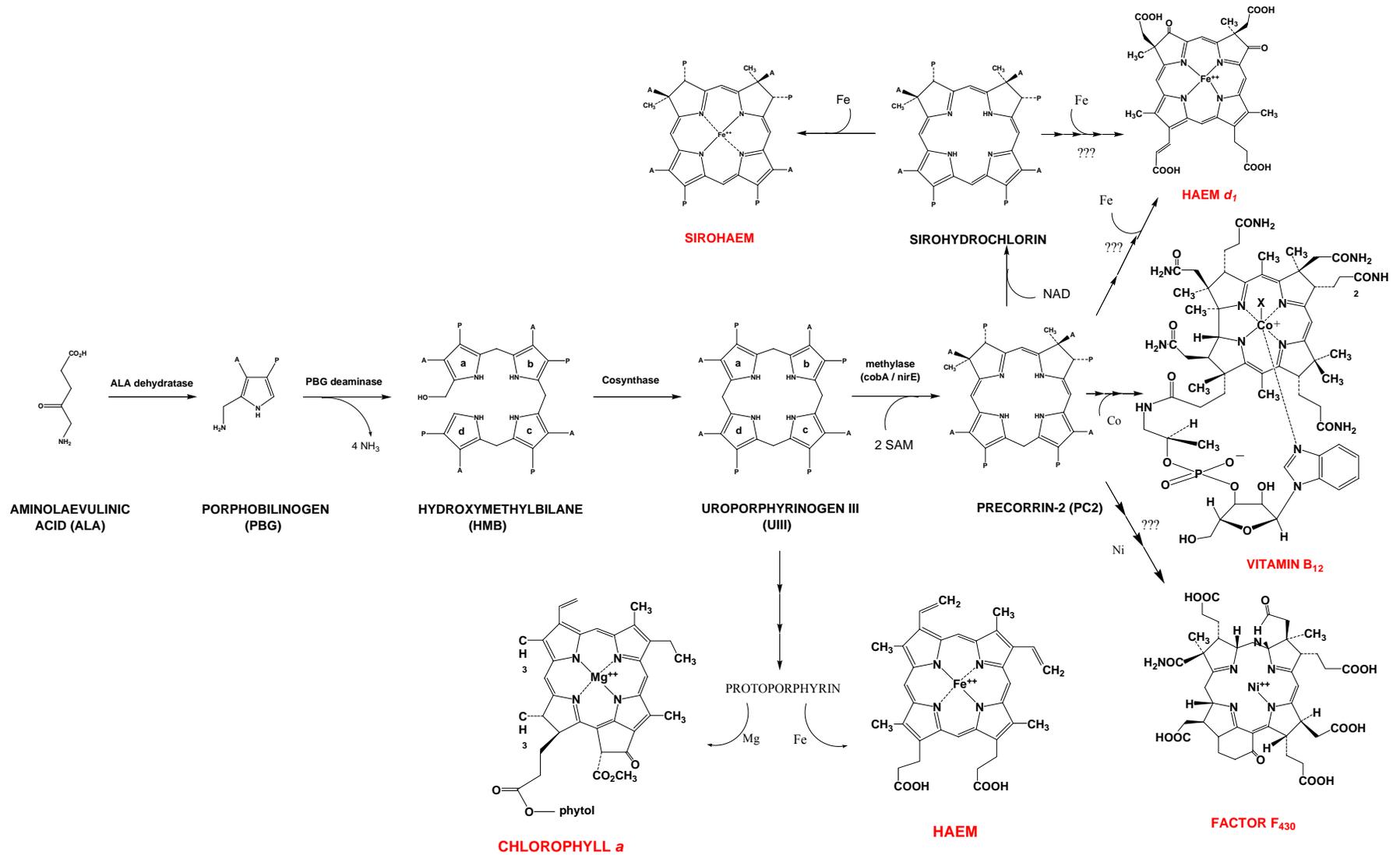


Figure 1.2: Summary of Tetrapyrrole biosynthesis: a branched pathway from Uroporphyrinogen III (UIII).

### 1.1.1. Chlorophyll and Bacteriochlorophylls

Chlorophylls and bacteriochlorophylls are characteristically green (or purple) compounds that are essential for the conversion of solar energy to chemical energy as part of photosynthesis. They all share the same macrocyclic ring, possessing an additional fifth ring (ring E) formed by cyclisation of a propionic acid side chain and a centrally chelated  $Mg^{2+}$ .

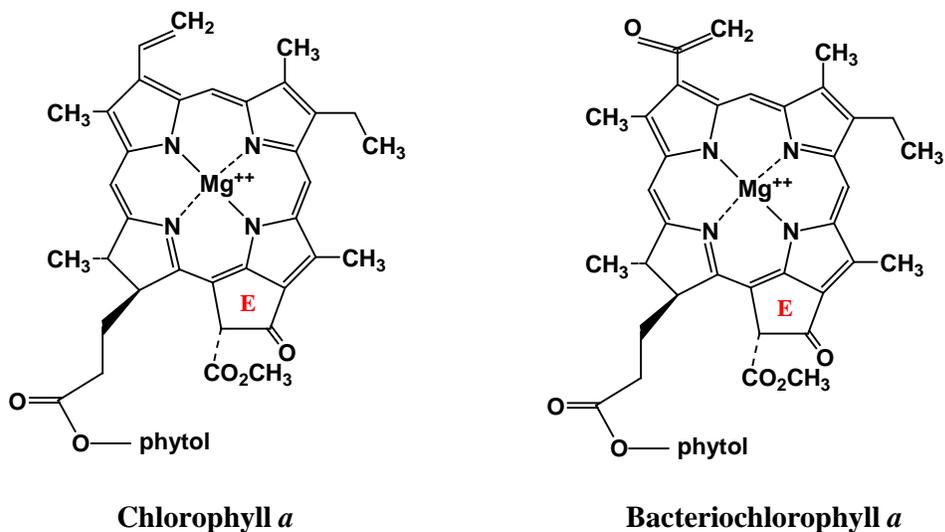
Chlorophylls can be divided into three primary classes of chlorophylls, *a*, *b* and *c*. Higher plants and green algae contain a small proportion of chlorophyll *b* but mostly chlorophyll *a*, the most ubiquitous and well characterized form (Figure 1.3). Chlorophyll *c* is found in eukaryotic algae but do not contain chlorophyll *b* (Beale 1990).

Bacteriochlorophylls are synthesised by purple and green sulfur bacteria of which there are six classes, *a*, *b*, *c*, *d*, *e* and *g*. Bacteriochlorophyll *a*, the most common (Figure 1.3), and *b* are purple pigments produced by photosynthetic bacteria (e.g. *Rhodobacter sphaeroides* (Bchl *a*)). Bacteriochlorophyll *c*, *d* and *e* are pigments of green and brown sulfur bacteria (e.g. *Prosthecochloris aestuarii* (BChl *c*)), and bacteriochlorophyll *g* has been found in *Heliobacterium chlorum* (Smith 1991).

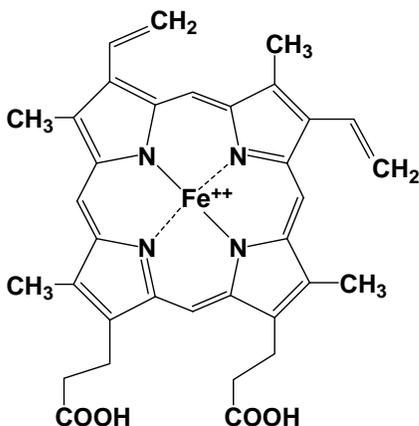
The biosynthesis of all chlorophylls and bacteriochlorophylls is a shared pathway of which chlorophyll *a* is well studied. It has been shown to share the same pathway as haem (Figure 1.5), up to the formation of protoporphyrin IX (PPIX), after which  $Mg^{2+}$  is inserted into PPIX by a trimeric ATP-dependent enzyme, ChlH, ChlD and ChlI (or BchH, BchD and BchI for bacteriochlorophyll synthesis). BchE catalyses the addition of the extra ring (ring E), and the phytol group is subsequently attached (Beale 1990).

**Figure 1.3: The structures of Chlorophyll *a* and Bacteriochlorophyll *a***

Chlorophyll *a* and bacteriochlorophyll *a*, are green in colour and are depicted below along with the additional ring, **E**. The formation of each from UIII has required decarboxylation at C2, C7, C12 and C18 to produce methyl groups, whilst C3 and C8 have undergone oxidation to form vinyl and alkyl groups respectively. Chelation of  $Mg^{2+}$  occurs, and finally phytol group is attached to the propionate group at C17.

**Figure 1.4: The structure of Haem *b***

Haem *b* is red in colour giving haemoglobin its distinctive pigmentation; the entire pathway from UIII to haem *b* is shown in Figure 1.5.



### 1.1.2. Haem

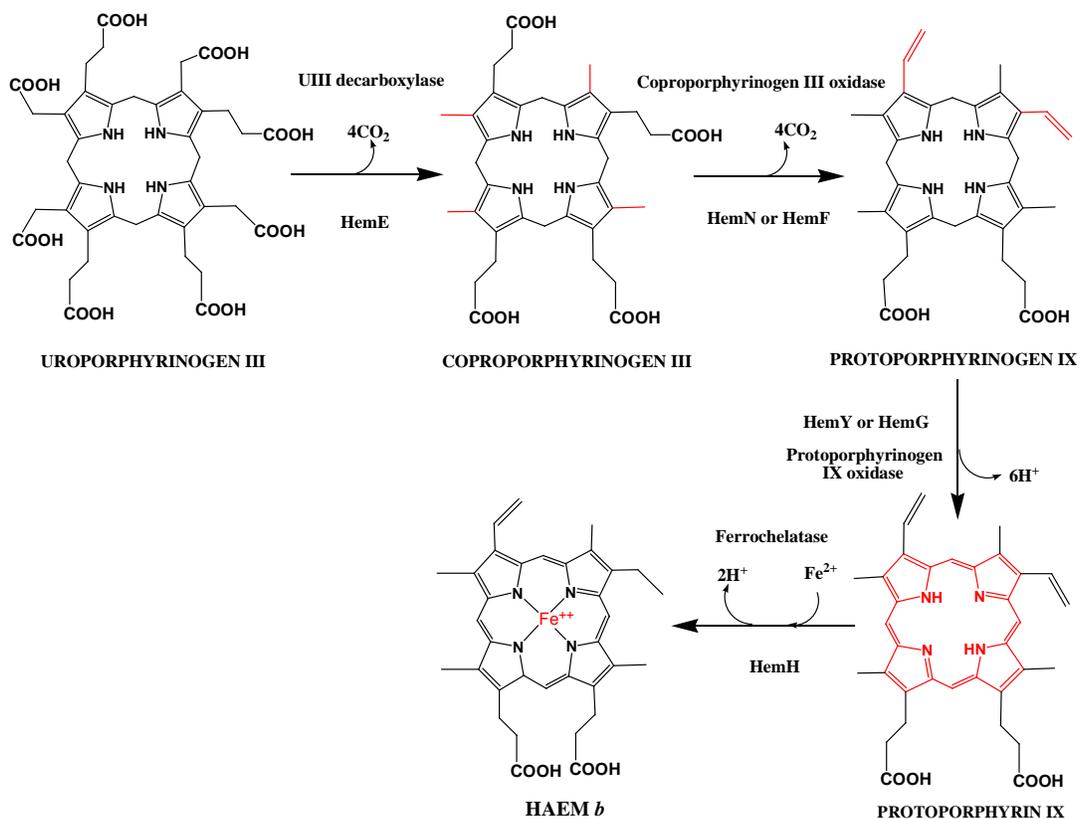
One of the most colourful molecules found in nature is the haem molecule. Each exhibits unique properties based on the modifications of its peripheral groups with a ferrous ion at its core. Five classes of haem, differing only in their side chains, are synthesised in nature: *a*, *b*, *c*, *d* and *o*. Each type of haem acts as the prosthetic group of vital enzymes and proteins required in fundamental and regulatory processes nature. Typically, haems *a*, *b* and *c* are red whilst haems *d* and *o* are green. (Chang 1994)

Haem *a* is found in cytochrome oxidases catalysing the terminal reaction of the respiratory chain. The *b*-type haems (Figure 1.4), are the most well known and include protohaem as found in haemoglobin, cytochromes *b* and  $P_{450}$ . Haem *c* is found in cytochrome *c* proteins. Haem *d* and *o* are associated with terminal cytochrome oxidases mediating the reduction of oxygen to water (Beale 1990; Chang 1994). Haem *d* and Haem  $d_1$  exhibit unique properties not shared by *b*-type haems due to the addition of oxygen to the tetrapyrrole macrocycle thus altering their  $\pi$ -orbital system. Of these haem  $d_1$  mediates the reduction of nitrite to nitric oxide as well as the four electron reduction of oxygen to water (Silvestrini, Falcinelli et al. 1994). Haem  $d_1$  will be discussed in greater detail later.

Haem biosynthesis is mediated under both aerobic and anaerobic conditions, of which the aerobic pathway has been more extensively characterized, as outlined in Figure 1.5. The anaerobic pathways involved in haem biosynthesis are much less well characterized as in the case of haem  $d_1$ . An alternative haem pathway has been shown to exist in the sulfate reducing anaerobe *Desulfovibrio vulgaris*. In this bacterium protohaem (haem *b*), is synthesised by the methylation of UIII (C2, C7) to form PC-2, which is then decarboxylated (C12, C18) and dealkylated (C2, C7), yielding coproporphyrinogen III, an intermediate in the haem *b* biosynthetic pathway (Ishida, Yu et al. 1998). Thus, there is evidence for the employment of different pathways that share similar intermediates but employ different enzymes.

**Figure 1.5: The haem *b* biosynthetic pathway**

The formation of this modified tetrapyrrole from UIII requires decarboxylation of the acetic acid side chains to produce methyl groups at C2, C7, C12 and C18, by UIII decarboxylase yielding coproporphyrinogen III, and the release of four  $\text{CO}_2$  molecules. Subsequently, the propionate groups at C3 and C8 are reduced to vinyl groups by coproporphyrinogen III oxidase to give protoporphyrinogen IX and the release of two  $\text{CO}_2$  molecules. The macrocyclic ring undergoes a six-electron oxidation to form protoporphyrin IX, inducing a change in the entire ring structure as shown in red. Up to this point the pathway for haem and chlorophyll are undifferentiated, chelation of  $\text{Mg}^{2+}$  directs flux towards the production of chlorophyll whereas chelation of  $\text{Fe}^{2+}$  completes haem biosynthesis.



### 1.1.3. Sirohaem

Sirohaem (Figure 1.6) is an isobacteriochlorin. It is the prosthetic group of nitrite and sulfite reductases, which convert nitrite to ammonia and sulfite to sulfide respectively, in a six-electron reduction step. Both these reductases depend on the sirohaem moiety to assist in this reaction (Murphy, Siegel et al. 1974; Crane, Siegel et al. 1995). The absence of sirohaem from sulfite reductase therefore leads to the failure to reduce sulfite and phenotypically presents as a cysteine auxotrophy since it is sulfide that reacts with *O*-acetylserine to form cysteine (Figure 1.7) (Kredich 1987).

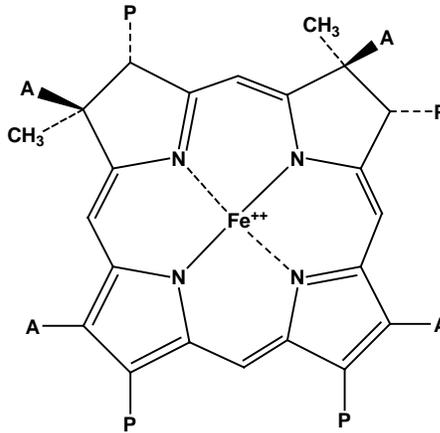
Sirohaem, being a much less complicated cofactor, requires significantly fewer steps for synthesis than does cobalamin. From UIII, two SAM dependent methylations generate PC-2, via PC-1 (Woodcock and Warren 1996), by methylation at C2 then C7. PC-2 is then dehydrogenated in a nicotinamide adenine dinucleotide (NAD<sup>+</sup>) dependent manner to give sirohydrochlorin (SHC), which is then chelated with a ferrous ion to complete the synthesis of sirohaem.

For a simple pathway, there are many similar enzymes that catalyse the formation of sirohaem, as summarized in Figure 1.8. In *B. subtilis* and *B. megaterium*, three monofunctional enzymes, SirA, SirC and SirB, catalyse the methylation, reduction and ferrochelation, respectively (Leech, Raux-Deery et al. 2002; Raux, Leech et al. 2003). In yeast, *S. cerevisiae* the transformation of UIII into sirohaem is performed by only two enzymes; The SAM-dependent methylation of UIII is catalysed by a single enzyme (Met1p) but dehydrogenation and chelation steps are catalysed by a bi-functional enzyme, Met8p (Hansen, Muldbjerg et al. 1997; Raux, McVeigh et al. 1999). It is notable that in some bacteria, only a single enzyme performs the conversion of UIII into sirohaem. In *S. enterica* and in *E. coli* a tri-functional enzyme, CysG, is utilized. Studies of CysG reveal that it is composed of 2 distinct catalytic domains, CysG<sup>A</sup> and CysG<sup>B</sup> (Spencer, Stolowich et al. 1993; Warren, Bolt et al. 1994; Stroupe, Leech et al. 2003). CysG<sup>A</sup> confers SAM dependent methylation activity transforming UIII to PC-2 and has high levels of sequence similarity to MET1p and SirA, which function independently (Warren, Bolt et al. 1994; Raux, McVeigh et al. 1999). CysG<sup>B</sup> is comprised of the catalytic region responsible for NAD<sup>+</sup> dependent dehydrogenation and ferrochelation. MET8p and SirC also display similarity to CysG<sup>B</sup>, but SirB has no sequence identity with

CysG<sup>B</sup>, suggesting CysG is somewhat unique in that the gene fusion event that created this trifunctional enzyme (Warren, Roessner et al. 1990; Warren, Bolt et al. 1994).

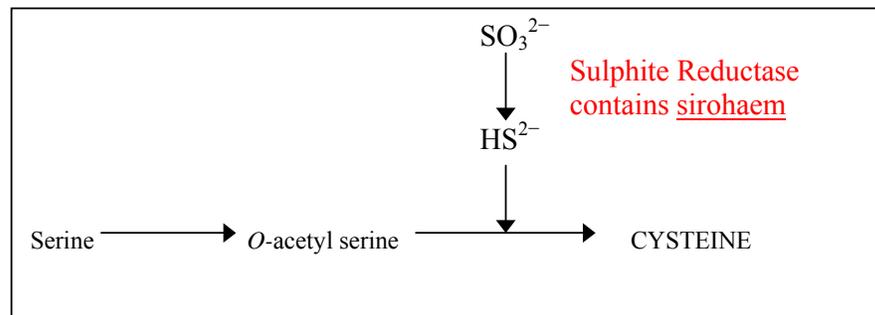
**Figure 1.6: The structure of Sirohaem**

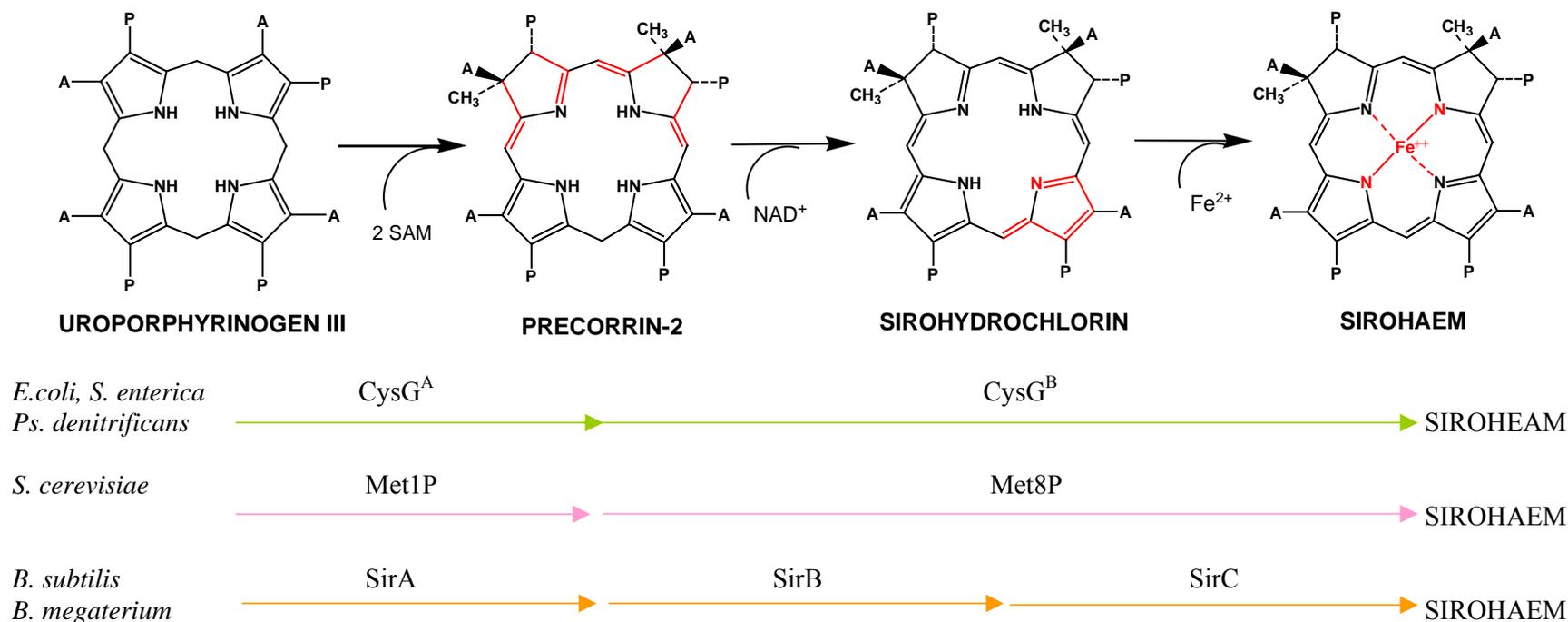
A = CH<sub>2</sub>CO<sub>2</sub>H, P = CH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H.



**Figure 1.7: The role of Sirohaem in cysteine biosynthesis**

The biosynthesis of cysteine from serine requires sulphide, produced by the sirohaem containing sulphite reductase, since it is sulphide that reacts with *O*-acetylserine to form cysteine. The absence of sirohaem from sulphite reductase leads to failure to reduce sulphite and phenotypically presents as a cysteine auxotrophy





**Figure 1.8: Diversity in the biosynthetic pathway of the transformation of Uroporphyrinogen III into Sirohaem in different organisms**

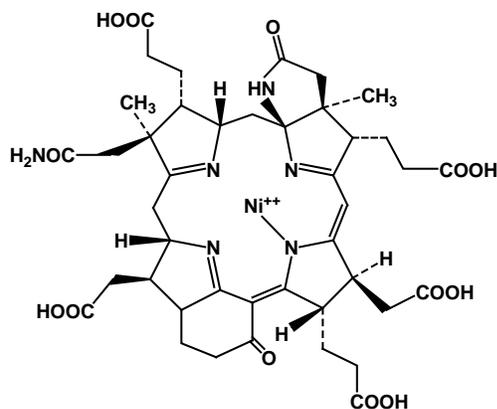
Depending on the organism, the conversion of UIII to Sirohaem is performed by different enzymes. *E. coli* and *S. enterica* generate sirohaem using a single multifunctional enzyme (CysG) with 2 domains (CysG<sup>A</sup> and CysG<sup>B</sup>), shown in green. *S. cerevisiae* utilises a 2-step process (pink). *B. megaterium* and *B. subtilis* utilises a 3-step pathway, with SirA, SirB, SirC, shown in orange. The changes in each structure along the pathway are highlighted in red. A = CH<sub>2</sub>CO<sub>2</sub>H P = CH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H.

### 1.1.4. Coenzyme F<sub>430</sub>

Coenzyme F<sub>430</sub> is the prosthetic group of methyl coenzyme M reductase, which catalyses the final step of methane formation unique to methanogenic bacteria such as *Methanobacterium thermoautotrophicum*. It is a nickel-containing yellow-pigmented modified tetrapyrrole (Figure 1.9). The crystal structure of methyl CoM reductase has been resolved to 1.45 Å, and shown to contain two molecules of coenzyme F<sub>430</sub>. The role of coenzyme F<sub>430</sub> in the formation of global methane has been well described (Ermler, Grabarse et al. 1997; Shima, Goubeaud et al. 1997). Despite the important role that coenzyme F<sub>430</sub> plays in nature, very little is known about the synthesis of this nickel containing tetrapyrrole. Other previous work suggests that it is synthesised from SHC and one intermediate, 15,17(3)-seco-F430-17(3)-acid has been reported (Pfaltz, Kobelt et al. 1987).

#### Figure 1.9: The structure of Coenzyme F<sub>430</sub>

The synthesis of this molecule remains to be solved. It has two additional rings attached to the tetrapyrrole macrocycle, mediated by closure of the side chains. A Ni ion is centrally chelated to the molecule.



### 1.1.5. Vitamin B<sub>12</sub>

Vitamin B<sub>12</sub> is one of the most important and complicated of tetrapyrroles found in nature. The resolution of the x-ray crystal structure of cyanocobalamin by Dorothy Hodgkin revealed its great complexity (Hodgkin, Pickworth et al. 1955). A cobalt ion is centrally ligated to the contracted macrocycle of the tetrapyrrole (rings A and D are bonded together), the cobalt is further bound to a modified base on the lower axial ligand (a dimethyl benzimidazole) and also an upper axial ligand, shown as X in Figure 1.10.

The upper axial ligand imposes different functions to the vitamin B<sub>12</sub> molecule. The presence of an upper axial cyano group is an artefact of the extraction procedure. In biological systems there are two principle forms, adenosyl cobalamin and methyl cobalamin, in which the former is used to catalyse rearrangement or reductase reactions and the latter methyltransferase reactions (Warren, Raux et al. 2002).

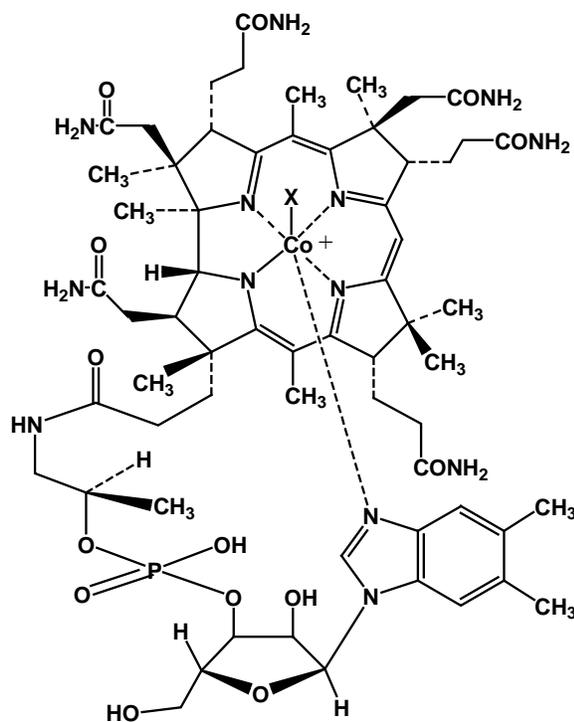
The complexity in its structure is also indicative of its intricate requirement for approximately thirty genes for the *de novo* synthesis of this molecule. Vitamin B<sub>12</sub> synthesis is uniquely restricted to few organisms, and no known eukaryote possesses the ability to make cobalamin. However, eukaryotes that require cobalamin rely on eubacteria and archeobacteria as sources and some have evolved to have intricate uptake systems (Seetharam and Alpers 1982).

There are two distinct alternative routes towards the synthesis of vitamin B<sub>12</sub> in bacteria and archaea. These are differentiated by the requirement of oxygen and the absence of oxygen, termed the aerobic and anaerobic pathways. The pathway from ALA to PC-2 is the same in both routes. The aerobic pathway incorporates molecular oxygen in to the macrocycle as a prerequisite to ring contraction and cobalt is inserted late in the pathway. The aerobic route has been extensively studied in *Ps. denitrificans*, with the isolation of intermediates and the genes required (Debussche, Thibaut et al. 1993; Blanche 1995). The anaerobic route is characterised by the early insertion of cobalt at the PC-2 stage and the oxygen independent ring contraction. The anaerobic route is the more ancient to its aerobic counterpart (Santander, Roessner et al. 1997) and has been studied in *S. enterica* (Raux, Lanois et al. 1996), *B. megatarium* (Raux, Lanois et al. 1998; Raux, Lanois et al. 1998; Raux, Schubert et al. 1998) but is incomplete. Genes from organisms that make

cobalamin aerobically are referred to as *cob* and those that make cobalamin anaerobically are denoted the prefix *cbi*.

**Figure 1.10: The structure of Vitamin B<sub>12</sub>**

The structure of this molecule is of a complex nature. Depending upon the function, X can be a methyl group, a cyano group or a 5'-deoxyadenosine, bestowing the macrocycle with different properties.



## 1.2. Haem $d_1$

Haem  $d_1$  is a modified tetrapyrrole with a characteristic green colour and a centrally chelated iron moiety. It has many unique features that have made it a source of great interest for over 20 years.

It is uniquely found as the prosthetic group of periplasmic bacterial dissimilatory nitrite reductase, also known as cytochrome  $cd_1$ , catalysing the dissimilatory reduction of nitrite ( $\text{NO}_2^-$ ) to nitric oxide (NO), as part of denitrification. This occurs in all denitrifying bacteria during anaerobic respiration. Thus, haem  $d_1$  is also synthesised under anaerobic conditions. On a global scale, this is a crucial step in all denitrifying bacteria as part of the global nitrogen cycle, where denitrification is the committing step responsible for the return of fixed terrestrial nitrogen to the atmosphere on a substantial scale (Averill 1996; Zumft 1997).

The structure of haem  $d_1$  has posed more questions in recent history than it answers. The green colour of haem  $d_1$  led many to assume that it possessed a chlorin core structure. In 1984, Timkovich *et al* first proposed a porphyrindione (dioxoisobacteriochlorin) structure (Timkovich, Cork *et al.* 1984; Timkovich, Cork *et al.* 1984a) and after much debate, spectroscopic analysis and chemical synthesis (Chang 1985; Chang and Wu 1986), the structure was proven to be that of a porphyrindione (Chang, Timkovich *et al.* 1986b).

Haem  $d_1$  differs from other haems and tetrapyrrole derivatives in that its synthesis is mediated via a completely separate branch of the tetrapyrrole biosynthetic pathway from UIII (Figure 1.2). Furthermore, it is a dioxoisobacteriochlorin, as opposed to a porphyrin, characterised by the presence of two oxo groups at C3 and C8, methyl groups at C2 and C7 and an acrylate in the C17 peripheral group as depicted in Figure 1.11.

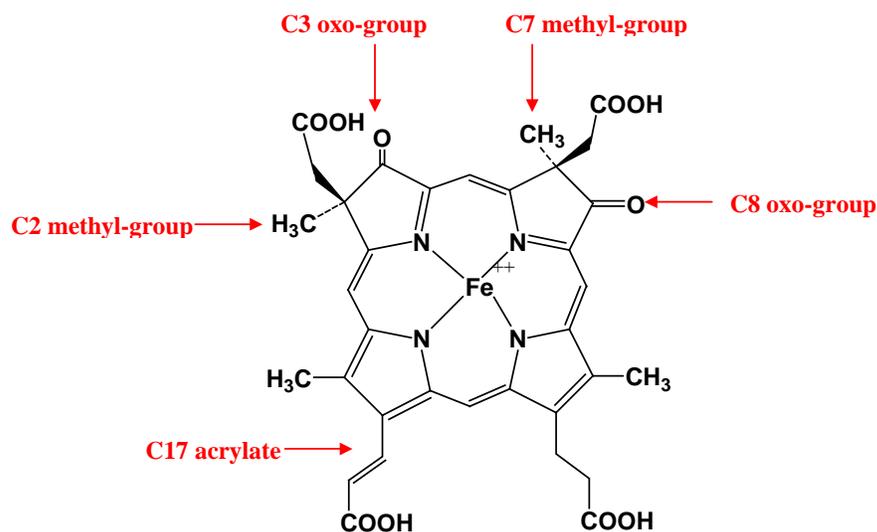
In comparison to haem  $b$ , haem  $d_1$  differs in that two pyrrole units are more saturated and the presence of two keto (oxo) groups at C3 and C8 alters the  $\pi$ -orbital configuration around the macrocycle; further saturation is observed with an acrylate in the C17 peripheral group. Such features give haem  $d_1$  unique properties not shared by other haems. In particular, the presence of the oxo groups acts to pull down the  $\pi$ -stack to bring the redox potential of haem  $d_1$  closer to the parent porphyrin and thus

more positive than those of other porphyrin haems (Chang, Hanson et al. 1981). It has recently been resolved to be + 175mV in *Paracoccus pantotrophus*, by way of shifting the  $c$ -haem to - 60mV (Zajicek, Cartron et al. 2006). This redox property is advantageous for the ease of substrate reduction that haem  $d_1$  partakes in (Section 1.2.1). Furthermore, the oxo groups make the macrocycle harder to oxidise.

A link to the tetrapyrrole pathway was established by the stereoselective synthesis of haem  $d_1$  which showed the absolute configuration at C2 and C7 matched the 2R, 7R configuration also seen in PC-2, sirohydrochlorin (SHC), and vitamin B<sub>12</sub> (Micklefield 1993; Micklefield 1997). The introduction of [<sup>13</sup>C-*methyl*] methionine to the growth medium of *Ps. aeruginosa* showed that the methyl groups at C2 and C7 are derived from *S*-adenosyl-L-methionine (SAM) (Yap-Bondoc, Bondoc et al. 1990). This is in a similar fashion to vitamin B<sub>12</sub>, sirohaem and coenzyme F<sub>430</sub>, thereby connecting haem  $d_1$  synthesis to UIII and PC-2 in the tetrapyrrole pathway.

**Figure 1.11: The structure of Haem  $d_1$**

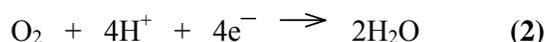
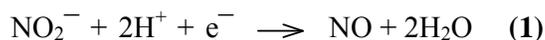
The structure of haem  $d_1$  is characterised by the presence of two oxo groups at C3 and C8, methyl groups at C2 and C7, and an acrylate in the C17 peripheral group and a centrally chelated iron.



### 1.2.1. Haem $d_1$ in cytochrome $cd_1$

Haem  $d_1$  is the prosthetic group of periplasmic bacterial dissimilatory nitrite reductase or cytochrome  $cd_1$ . It has also previously been referred to as cytochrome oxidase due to its catalysis of oxygen to water (Horio, Higashi et al. 1961; Silvestrini, Falcinelli et al. 1994).

Cytochrome  $cd_1$  is a soluble homodimeric enzyme of 120 kDa. It is a bifunctional enzyme that catalyses both the one electron reduction of nitrite to nitric oxide (1) and the *in vitro* four electron reduction of oxygen to water (2)



As the name suggests, each equivalent monomer of 60kDa contains a covalently bound *c*-type haem via two Cys residues, and a non-covalently bound haem  $d_1$  (Kuronen and Ellfolk 1972; Gudat, Singh et al. 1973; Averill 1996). There exists a wealth of experimental evidence suggesting that the haem  $d_1$  iron moiety provides the active centre for nitrite reduction and oxygen reduction, and haem *c* is the site of electron entry from donors such as azurin, pseudoazurin and cytochrome  $c_{551}$  (Silvestrini, Tordi et al. 1990; Fulop, Moir et al. 1995; Baker, Saunders et al. 1997; Williams, Fulop et al. 1997; George, Allen et al. 2000; Pearson, Page et al. 2003; Ferguson 2009).

The 3D crystal structures of cytochrome  $cd_1$  have been determined from *Pa. pantotrophus* (previously *Thiosphaera pantotropha*) (Fulop, Moir et al. 1995), (Figure 1.12), and *Ps. aeruginosa* (Nurizzo, Silvestrini et al. 1997). The monomers consist of two structural domains: an N-terminal  $\alpha$ -helical *c*-domain and a C-terminal  $\beta$ -propeller  $d_1$  domain, where the  $d_1$  haem is held within an eight bladed  $\beta$ -propeller structure (Figure 1.13) (Baker, Saunders et al. 1997). In spite of similar architecture, the structure of the haem binding sites differs significantly in the two enzymes from each organism (Cheesman, Ferguson et al. 1997). In the oxidised *Pa. pantotropha* cytochrome  $cd_1$ , His17 and His69 provide the axial iron ligands for haem *c*; His200 is a fifth axial iron ligand for haem  $d_1$ . An unusual feature of cytochrome  $cd_1$  is that the main chain of the *c* domain extrudes into the  $d_1$  domain, providing Tyr25 as the sixth axial iron ligand for haem  $d_1$  (Fulop, Moir et al. 1995). Reduction, allows Tyr25 to be released and allow substrate binding to haem  $d_1$ ;

concomitantly, a refolding of the  $c$  domain takes place, resulting in a switch of one haem  $c$  iron ligand from His17 to Met106 (Williams, Fulop et al. 1997). In the oxidised *Ps. aeruginosa* and *Ps. stutzeri* protein, His51 and Met88 provide the axial iron ligands for haem  $c$  and His182 provides the fifth axial iron ligand for haem  $d_1$ . The hydroxide ion, hydrogen bonded to Tyr10, is a sixth axial iron ligand for haem  $d_1$  (Cheesman, Ferguson et al. 1997). Interestingly, no ligand exchange has been observed in *Ps. aeruginosa* cytochrome  $cd_1$ , but a conformational change has been observed on binding of the substrate (Nurizzo, Silvestrini et al. 1997; Nurizzo, Cutruzzola et al. 1998; Nurizzo, Cutruzzola et al. 1999). Mutational analysis has implicated the requirement of His327 and His369 in the activity of the enzyme (Cutruzzola, Brown et al. 2001).

Thus, although both sources of cytochrome  $cd_1$  share a similar function they have different haem ligands. Clearly, the presence of a tetrapyrrole as intricate the  $d_1$  haem within this single enzyme with different ligands has evolved for a specific purpose and poses a catalytic advantage in the reduction of nitrite to nitric oxide.

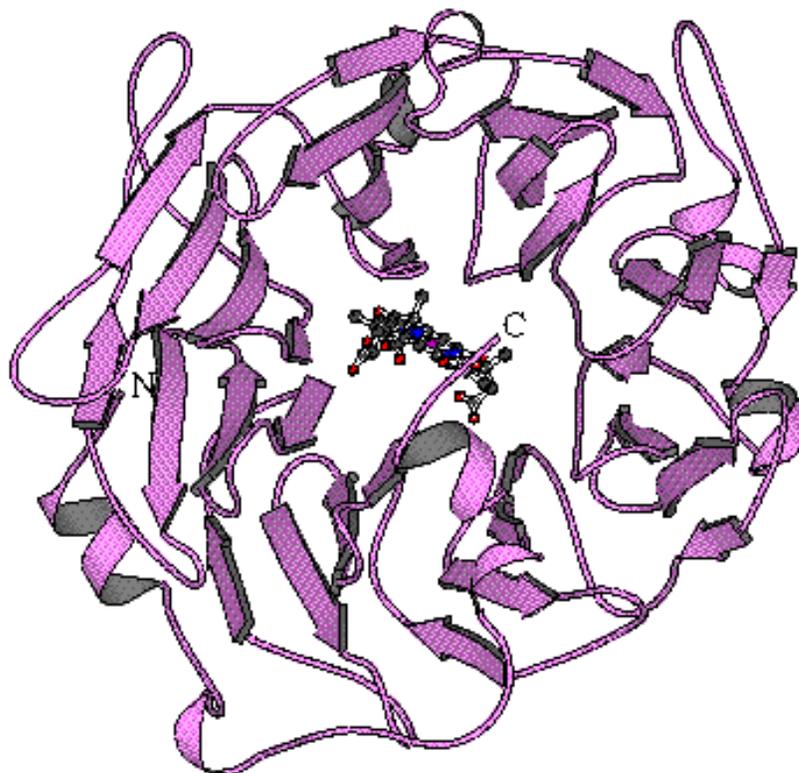
### Figure 1.12 Cytochrome $cd_1$ structure

Ribbon representation of *Pa. pantotrophus* cytochrome  $cd_1$  homodimeric structure. Colour codes: chain A = orange; chain B = violet. The haem  $c$  and haem  $d_1$  domains with the bound structures are indicated. Adapted from (Fulop, Moir et al. 1995).



**Figure 1.13:  $\beta$ -propeller  $d_1$  domain structure**

Cytochrome  $cd_1$   $\beta$ -propeller  $d_1$  domain (residues 135-624) shows the location of the  $d_1$  haem held non-covalently within the centre of the pocket. It is an eight bladed  $\beta$ -propeller structure. Adapted from (Fulop, Moir et al. 1995).



### 1.2.2. Denitrification

Denitrification is an important chemical transformation in the global nitrogen cycle converting nitrate,  $\text{NO}_3^-$ , to nitrogen gas,  $\text{N}_2$ , in four sequential steps. It is the only process substantially returning fixed terrestrial nitrogen to the atmosphere being sustained by bacteria, to complete the nitrogen cycle. Moreover, denitrification contributes to the greenhouse effect (Lashof 1990; Hernandez-Ramirez, Brouder et al. 2009) and has been linked to the destruction of the ozone layer (Waibel, Peter et al. 1999), both being key environmental issues.

Denitrification is the anaerobic use by facultative bacteria, of nitrogen oxide species as alternative terminal electron acceptors during respiration in the place of  $\text{O}_2$ , for the generation of the electrochemical gradient across the cytoplasmic membrane (Ye, Averill et al. 1994; Berks, Ferguson et al. 1995b). The most detailed chromosomal map of a denitrifying bacterium is that of *Ps. aeruginosa* (Holloway, Romling et al. 1994), of which those for anaerobic respiration and denitrification are mapped (Vollack, Xie et al. 1998). Much work has focused on the identification of denitrifying genes in *Ps. aeruginosa*, *Ps. stutzeri*, *Ps. fluorescens* and *Pa. denitrificans* and has been reviewed (Averill 1996; Zumft 1997; Baker, Ferguson et al. 1998; Philippot 2002). Denitrification also extends beyond bacteria to archaea, and is found in halophilic and hyperthermophilic archaea (Philippot 2002).

The process of denitrification is the reduction of nitrate ( $\text{NO}_3^-$ ) to nitrite ( $\text{NO}_2^-$ ), catalysed by nitrate reductase;  $\text{NO}_2^-$  to nitric oxide (NO), by nitrite reductase; NO to nitrous oxide ( $\text{N}_2\text{O}$ ) by nitric oxide reductase, and then  $\text{N}_2$ , by nitrous oxide reductase. The genes encoding these functions for nitrate respiration (*nar*), nitrite respiration (*nir*), NO respiration (*nor*), and  $\text{N}_2\text{O}$  respiration (*nos*) are assembled in clusters in *Ps. stutzeri* (Jungst, Braun et al. 1991; Braun and Zumft 1992), *Ps. aeruginosa* (Arai, Igarashi et al. 1995) and *Pa. denitrificans* (de Boer, Reijnders et al. 1994; Berks, Page et al. 1995; de Boer, van der Oost et al. 1996). The reduction of the nitrogen oxide species (red) at each step is shown below; the enzymes (blue), each utilise a metal cofactor contributing in the reaction.



### 1.2.3. The denitrification enzyme system

The first step is catalysed by dissimilatory nitrate reductase (NaR), where a molybdenum cofactor provides the active site, to convert  $\text{NO}_3^-$  to  $\text{NO}_2^-$  (Hille 1996). The reduction of  $\text{NO}_2^-$  to NO is performed by dissimilatory nitrite reductase (NiR), where haem  $d_1$  is the active centre for the reaction. This is considered to be the committing step as it leads to a very reactive intermediate NO radical, possessing distinct chemical radical species properties that essentially are detrimental to the host, thus the process of denitrification has been committed. The presence of nitric oxide as a free intermediate has been a point of controversy, but the evidence for NO as an obligatory intermediate is now conclusive (Zumft 1993; Ye, Averill et al. 1994). Nitric oxide reductase (NoR) removes NO by converting it to  $\text{N}_2\text{O}$  with the aid of two *b* type haems. Finally,  $\text{N}_2$  is formed by the reduction of  $\text{N}_2\text{O}$  by nitrous oxide reductase (NoS), a multicopper enzyme.

Since these reactions are sequential it lends to the theory that they are performed in proximity to each other to facilitate sequential reactions with minimal toxicity. The organisation of these enzymes, particularly in gram-negative bacteria, such as *Pseudomonads* (facultative anaerobes) and *Paracoccus* (aerobe) has been well studied (Zumft 1997). The combined evidence as determined by cell fractionation, antibody labelling and electron microscopy (Coyne, Arunakumari et al. 1990; Korner and Mayer 1992; Zumft 1997) has been depicted in Figure 1.14. This model applies to *Ps. aeruginosa*, *Ps. stutzeri* and *Paracoccus* genus.

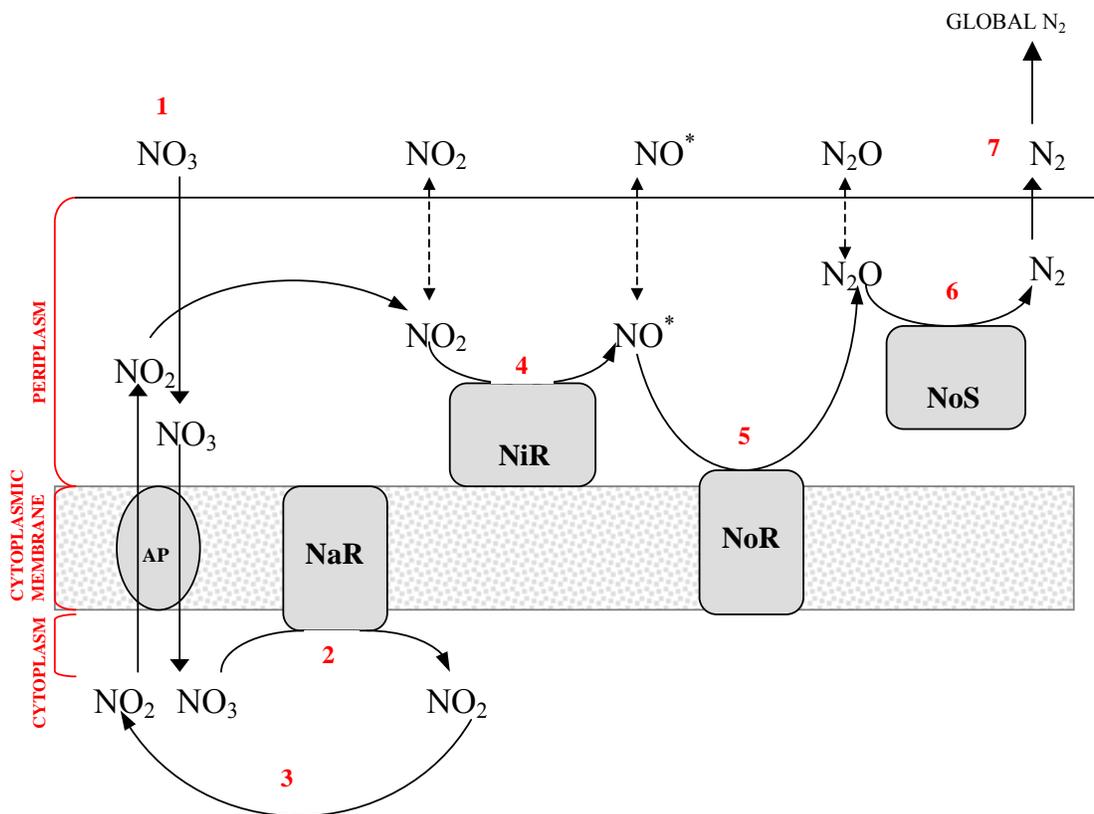
The first enzyme NaR, is situated in the cytoplasmic membrane with the active site exposed towards the cytoplasmic side (Carlson, Ferguson et al. 1982; Bell, Richardson et al. 1990), whereby the presence of NiR in the periplasmic space necessitates first the transport of  $\text{NO}_3^-$  across the periplasm and cytoplasmic membrane, to be reduced to  $\text{NO}_2^-$ ; Then its transportation to the periplasmic space which is postulated to be performed by a  $\text{NO}_3^-/\text{NO}_2^-$  antiporter.

NiR studied from *Ps. stutzeri* and *Ps. fluorescens* and NoS from *Ps. stutzeri*, have been shown to reside in the periplasmic space by immunochemical location (Coyne, Arunakumari et al. 1990; Korner and Mayer 1992; Zumft 1997). NiR in particular has been shown to be strongly associated with the cytoplasmic membrane of *Ps. aeruginosa* (Yamanaka 1992).

NoR has been localised to the inner membrane, whereby the reaction of NO to  $N_2O$  occurs on the periplasmic side (Bell, Richardson et al. 1990). This may have evolved as an evolutionary function so as to protect the cytoplasmic contents from radical /oxidative stress, as both NO and  $NO_2^-$  are toxic to the cell.

**Figure 1.14: Topology of the denitrification system in all denitrifying bacteria.**

Initially, the transport of  $NO_3^-$  through the periplasmic space and the cell membrane (1) is required for NaR, embedded in the membrane to reduce  $NO_3^-$  to  $NO_2^-$  on the cytoplasmic side (2). Nitrite then needs to be transported to the periplasmic space (3), probably via a  $NO_3^-/NO_2^-$  antiporter (AP), to be reduced to NO by NiR (4). NO, a species detrimental to the host, is converted to  $N_2O$  in the periplasm by the membrane enzyme NoR (5). Finally,  $N_2O$  is converted to  $N_2$  in the periplasm by NoS (6). Adapted from (Averill 1996).



### 1.3. The biosynthesis of Haem $d_1$

#### 1.3.1. Synthesis of 5-aminolaevulinic acid (ALA)

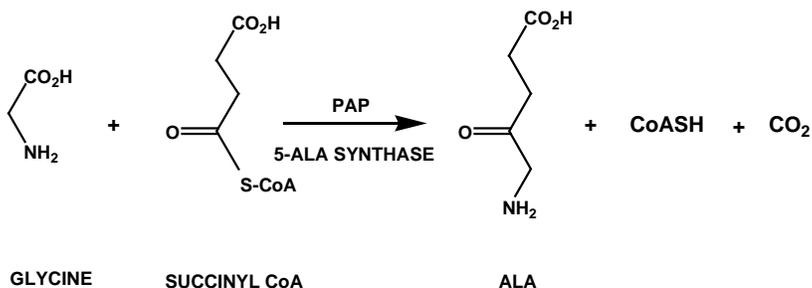
The first committed precursor in the biosynthesis of all tetrapyrroles is the synthesis of 5-aminolaevulinic acid (ALA). ALA is synthesised by one of two routes. In animals and photosynthetic bacteria (except higher plants and algae), succinyl coenzyme A and glycine condense to form ALA in a reaction catalysed by ALA synthase (HemA) with the loss of the carbonyl carbon of glycine (decarboxylation). This is known as the C-4 pathway, also called the Shemin route mediated by the *hemA* gene and requires pyridoxal phosphate as a co-factor (Figure 1.15). (Shemin 1946; Shemin 1989). Higher plants, algae, cyanobacteria and several other photosynthetic and non-photosynthetic bacteria and archaea produce ALA by the C-5 pathway, a multi-enzyme pathway that requires glutamyl-tRNA (tRNA<sup>Glu</sup>) (Schon, Krupp et al. 1986; Kannangara, Gough et al. 1988).

In *Ps. aeruginosa* and *Ps. stutzeri* the Shemin pathway is shown to not be operative in the synthesis of ALA but the C-5 pathway is employed (Matthews 1993; Hungerer, Troup et al. 1995a; Hungerer, Troup et al. 1995b). Here, ALA is synthesised from the intact carbon skeleton of glutamate by a multi-enzyme pathway. Binding tRNA<sup>Glu</sup> with ATP-dependant glutamyl-tRNA<sup>Glu</sup> synthetase activates Glutamate at the  $\alpha$ -carboxyl. An NADPH-dependent reductase (HemA) converts glutamyl-tRNA<sup>Glu</sup> to glutamate 1-semialdehyde, which is then transformed by glutamate 1-semialdehyde (GSA) aminotransferase (HemL) into ALA (Figure 1.16) (Huang and Wang 1986; Kannangara, Gough et al. 1988; Hennig, Grimm et al. 1997).

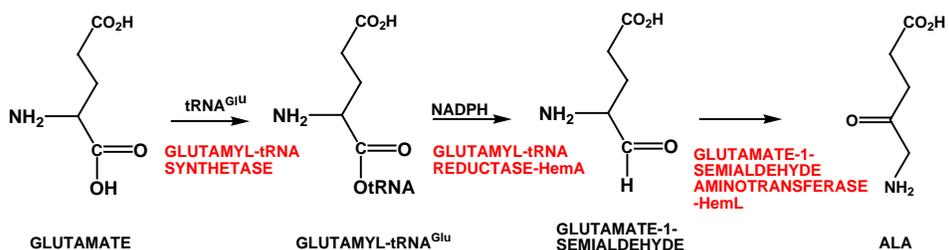
In *Ps. aeruginosa* and *Ps. stutzeri* the genes are mapped for *hemA* (Hungerer, Troup et al. 1995a) and *hemL* (Hungerer, Troup et al. 1995b), which are required for the synthesis of ALA. In *Pa. denitrificans* *hemA* has also been identified (Page and Ferguson 1994). Both *hemA* and *hemL* transcription is increased during denitrification in response to low oxygen and the presence of nitrate (Hungerer, Troup et al. 1995a; Hungerer, Troup et al. 1995b).

**Figure 1.15: Biosynthesis of ALA via the C-4 Shemin route**

Succinyl CoA and glycine condense to form ALA with a decarboxylation reaction. The reaction is catalysed by ALA synthase (HemA) and requires pyridoxal phosphate (PAP).

**Figure 1.16: Biosynthesis of ALA via the C-5 route**

ALA is formed from glutamate by this pathway. Glutamate binds tRNA<sup>Glu</sup> with glutamyl-tRNA<sup>Glu</sup> synthetase and NADPH-dependent reductase (HemA) converts glutamyl-tRNA<sup>Glu</sup> to glutamate 1-semialdehyde, which is then transformed by glutamate 1-semialdehyde (GSA) aminotransferase (HemL) into ALA.



### 1.3.2. Synthesis of the first macrocyclic intermediate, uroporphyrinogen III (UIII) from 5-aminolaevulinic acid (ALA)

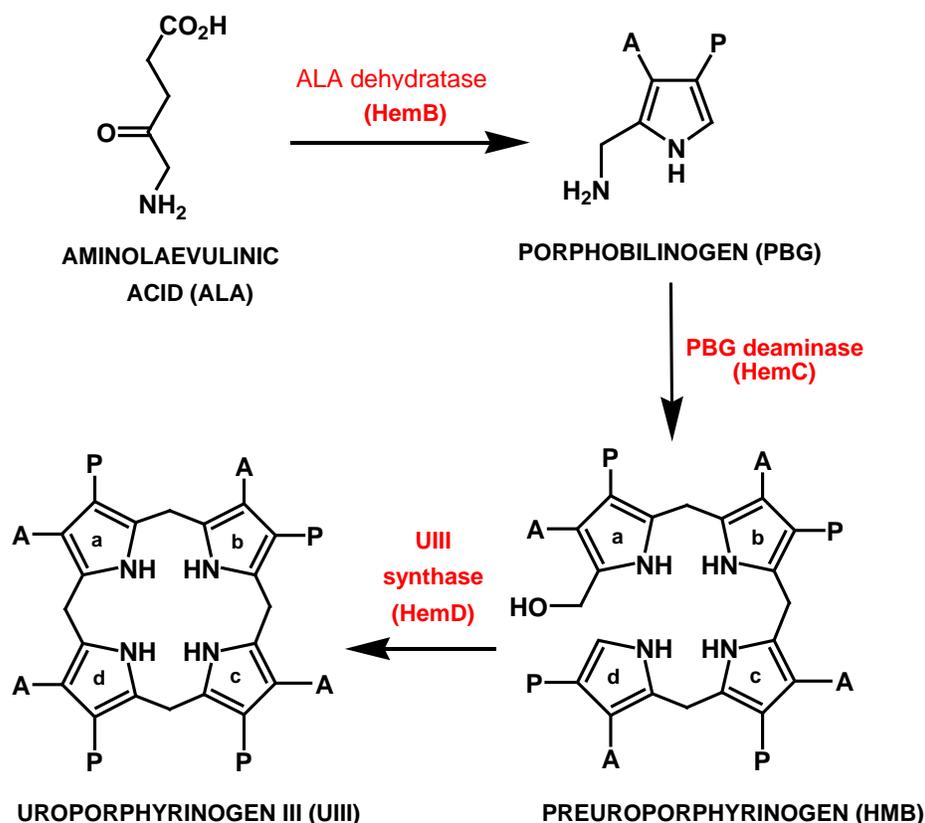
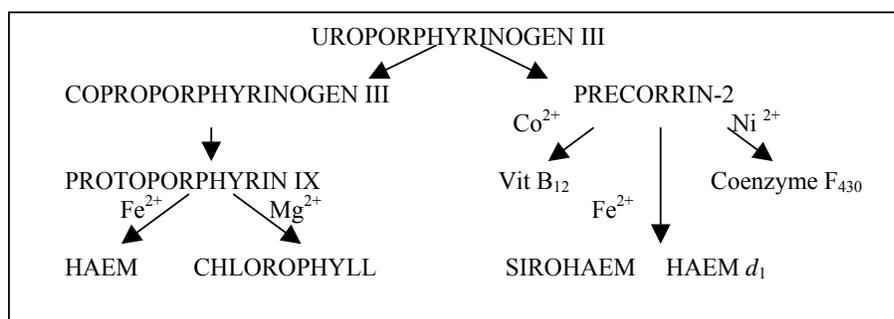
The transformation of ALA to UIII is common to all organisms. The genes encoding the enzymes required for this transformation have been discovered in a number of denitrifying bacteria as prerequisite to tetrapyrrole biosynthesis. The genes necessary are *hemB*, *hemC* and *hemD* that encode for the enzymes ALA dehydratase, porphobilinogen (PBG) deaminase and UIII synthase respectively (Figure 1.17). All *hemB*, *hemC* and *hemD* have been identified in *Ps. aeruginosa*.

Firstly, HemB (ALA dehydratase) catalyses the condensation of two molecules of ALA to form PBG, the first pyrrole structure. Next, HemC (PBG deaminase), catalyses the head to tail polymerisation of four molecules of PBG, leading to the generation of an open linear tetrapyrrole molecule named either hydroxymethylbilane (HMB) or preuroporphyrinogen. UIII is subsequently synthesised from preuroporphyrinogen by HemD (UIII synthase), which closes the tetrapyrrole to form the macrocyclic ring after inversion of ring D (Figure 1.17) (Warren and Scott 1990).

UIII is the last common intermediate from which the pathway branches to form the various modified tetrapyrroles. Each tetrapyrrole differs in its conjugation state, ring structure and chelated metal through the unique modification of the macrocycle from UIII. Enzymes specific to each pathway enable this and therefore bestow a wide range of structures and functions and unique properties. Decarboxylation of its acetate groups directs flux toward haem and chlorophyll synthesis whereas methylation of the macrocycle initiates synthesis toward a separate branch for sirohaem, vitamin B<sub>12</sub>, and haem  $d_1$  synthesis. A brief summary of the pathways deviating from UIII is shown in Figure 1.18.

**Figure 1.17: The biosynthetic conversion of ALA to UIII**

ALA is converted to UIII via 3 enzymatic steps. Two molecules of ALA condense to form PBG in a reaction catalysed by the Zn/Mg-dependent enzyme ALA dehydratase (ALAD) encoded for by *hemB*. Four molecules of PBG then polymerise forming the linear tetrapyrrole, pre-UIII or HMB (releasing ammonia). This reaction, catalysed by HemC, inverts the terminal ring forming the type III stereoisomer. The tetrapyrrole macrocycle is circularised by the closure of the tetrapyrrole ring, forming UIII. A =  $\text{CH}_2\text{CO}_2\text{H}$ , P =  $\text{CH}_2\text{CH}_2\text{CO}_2\text{H}$

**Figure 1.18: Summary of the tetrapyrrole pathway from UIII**

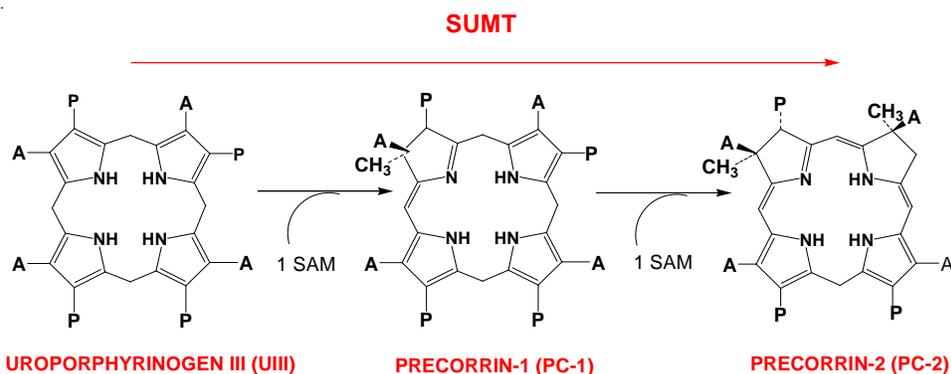
### 1.3.3. The methylation of uroporphyrinogen III (UIII) and the formation of precorrin-2 (PC-2)

Sirohaem, vitamin B<sub>12</sub>, and F<sub>430</sub> all require the methylation of UIII as part of their synthesis to form PC-2, also known as dihydrosirohydrochlorin. This reaction is catalysed by an enzyme specific to the process called *S*-adenosyl-L-methionine uroporphyrinogen III methyltransferase (SUMT). SUMT refers to a group of enzymes that have each evolved for a specific tetrapyrrole pathway but catalyse the same reaction.

As reflected in their name, all SUMT enzymes utilise *S*-adenosyl-L-methionine (SAM) as the methyl donor to insert two methyl groups onto the UIII macrocycle. This is executed by the subsequent methylation of UIII at the base of the acetate side chains at positions C2 and C7 to yield PC-2. This has been shown to occur via a mono-methylated intermediate Precorrin-1 concomitant with the release of the product *S*-adenosyl-L-homocysteine (SAH). Precorrin-1 is in turn methylated to produce the *bis*-methylated PC-2 (Figure 1.19) (Warren, Stolowich et al. 1990).

#### Figure 1.19: The transformation of UIII to PC-2 via PC-1

Analysis reveals that the transformation from UIII to PC-2 involves two SAM dependent methylations, firstly at C2 and then C7, catalysed by SUMT enzymes. This is achieved via a mono-methylated intermediate Precorrin-1 (PC-1) with the release of *S*-adenosyl-L-homocysteine (SAH), which is in turn methylated to produce the *bis*-methylated PC-2. A = CH<sub>2</sub>CO<sub>2</sub>H, P = CH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H.



There exist a plethora of SUMT enzymes that have been discovered in different organisms and plants that perform this transformation. The most well known and well studied is CobA.

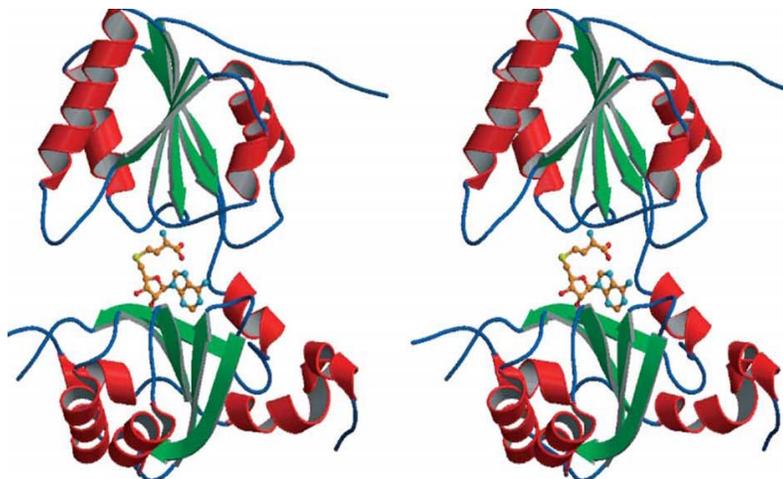
### 1.3.3.1. CobA in the vitamin B<sub>12</sub> pathway

The first SUMT to be characterised was CobA from *Ps. denitrificans* (Blanche, Debussche et al. 1989). It is a SUMT responsible for the C2 and C7 corrin ring methylations in the conversion of U<sub>III</sub> into PC-2, in the vitamin B<sub>12</sub> pathway. PC-2 represents the point at which the aerobic and anaerobic pathways diverge in their separate routes to cobalamin. Both pathways differ in the timing of metal insertion, method of the ring contraction and mechanism of carbon extrusion. The aerobic pathway is characterised by the requirement of molecular oxygen for the ring contraction (CobG) and subsequent release of acetic acid. The insertion of cobalt occurs late in this pathway and requires an ATP-dependent trimeric enzyme (CobNST). The anaerobic pathway, in contrast, inserts cobalt very early in the pathway (CbiK or CbiX) using considerably smaller, monomeric enzymes. It has no requirement for oxygen and less is known of the anaerobic pathway but is well reviewed (Warren, Raux et al. 2002; Frank, Brindley et al. 2005; Warren 2006).

*Ps. denitrificans* CobA is a homodimer with a subunit mass of 29kDa and a low catalytic efficiency (38/h) (Blanche, Debussche et al. 1989). The structure of this enzyme has been resolved to 2.7 Å with the product, SAH, still bound in the crystal (Figure 1.20) (Vevodova, Graham et al. 2004).

#### Figure 1.20: The structure of CobA from *Ps. denitrificans*

The homodimeric structure of CobA. Two domains are shown with the SAH ligand bound in the center of the proposed substrate-binding site. Each unit has five-stranded  $\beta$ -sheets with five associated  $\alpha$ -helices, and are made up of  $\beta$ - $\alpha$  repeating motifs. (Vevodova, Graham et al. 2004)



### 1.3.3.2. CysG<sup>A</sup> in *E. coli* and *S. enterica* in the sirohaem pathway

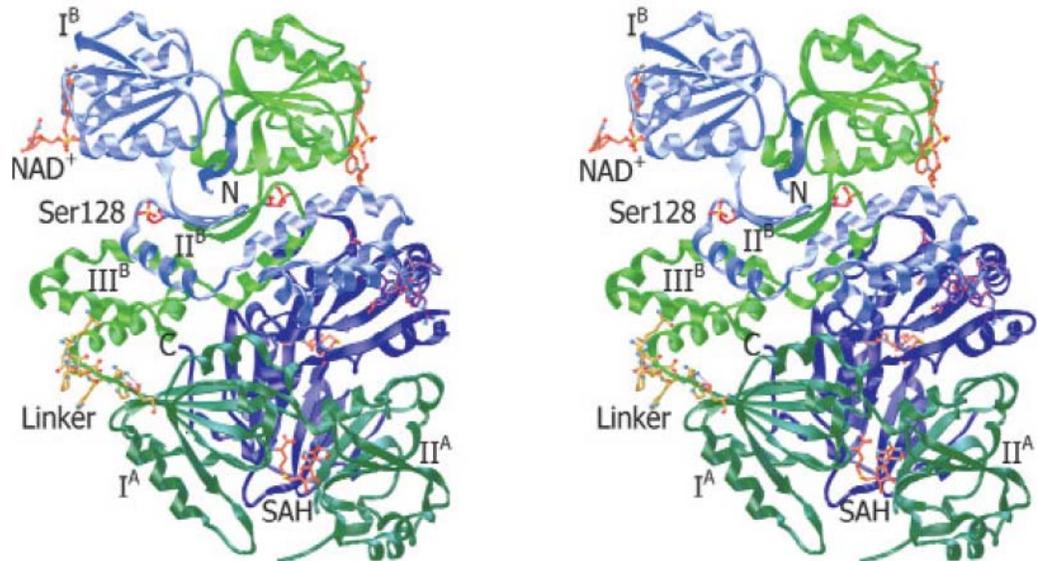
Sirohaem requires significantly fewer steps for synthesis than does cobalamin. From UIII, two SAM dependent methylations at C2 then C7 generate PC-2, via PC-1 (Woodcock and Warren 1996). PC-2 is then dehydrogenated in a NAD<sup>+</sup> dependent manner to give sirohydrochlorin (SHC), which is subsequently chelated with a ferrous ion to complete the synthesis of sirohaem. Sirohaem synthase or CysG is a unique multifunctional enzyme encompassing all these properties to transform UIII into sirohaem in *E. coli* and *S. enterica* (Peakman, Crouzet et al. 1990; Warren, Roessner et al. 1990; Wu, Siegel et al. 1991).

*E. coli* CysG is a 457 amino acid protein which can be described in terms of two functional domains; CysG<sup>A</sup> the C-terminal catalytic domain (amino acids 202-457-29KDa), which confers the transmethylase activity, and CysG<sup>B</sup>, the N-terminal catalytic domain (amino acids 1-201) which performs the dehydrogenation of SHC and ferrochelation reactions (Peakman, Crouzet et al. 1990; Warren, Roessner et al. 1990; Spencer, Stolowich et al. 1993; Warren, Bolt et al. 1994)

*E. coli* and *S. enterica* CysG have 90% similarity to each other at the primary structure level. Similar to CysG from *E. coli*, *S. enterica* CysG can be described as having two functional domains; CysG<sup>B</sup> (amino acids 1-214) conferring dehydrogenation and ferrochelate activity and CysG<sup>A</sup> that mediates transmethylase activity (amino acids 215-457) (Warren, Bolt et al. 1994). The structure of CysG from *S. enterica* has been resolved to 2.2 Å (Figure 1.21) (Stroupe, Leech et al. 2003). Resolving the structure of the CysG<sup>A</sup> domain revealed a resemblance to the structure of CbiF, a cobalt precorrin-4 methyltransferase enzyme from *B. megatarium* of the cobalamin pathway (Schubert, Wilson et al. 1998). Again, the structure was of a homodimer with two equivalent domains.

**Figure 1.21: The structure of CysG from *S. enterica***

The homodimeric structure of CysG is shown here from a  $\text{NAD}^+$  co-crystal structure. Each subunit is blue or green, where  $\text{CysG}^{\text{B}}$  at the top ( $\text{I}^{\text{B}}$ ) is more lightly coloured and  $\text{CysG}^{\text{A}}$  at the bottom ( $\text{I}^{\text{A}}$ ) is darkly coloured. The  $\text{NAD}$  and  $\text{SAH}$  are shown as orange. The linker regions from the blue and green subunits are purple and yellow respectively (Stroupe, Leech et al. 2003).



### 1.3.3.3. SirA in *Bacilli* in the sirohaem pathway

In *Bacilli* there are three separate enzymes responsible for the conversion of UIII into sirohaem, a methyltransferase, dehydrogenase and a chelatase. This is in contrast to the fused multifunctional CysG. (Leech, Raux-Deery et al. 2002) *B. subtilis* uses three proteins that are in a distinct gene cluster, YInD, the SUMT, YInF, the PC-2 dehydrogenase and YInE, the SHC ferrochelatase (Johansson and Hederstedt 1999).

In *B. megaterium*, the SUMT analogous to the *P. denitrificans* CobA was first identified in 1991 (Robin, Blanche et al. 1991). This SUMT of the sirohaem operon has since been renamed SirA and confirmed to be the UIII methyltransferase by mutation and complementation analysis as well as sequence analysis with other well-characterized SUMTs such as CobA and CysG<sup>A</sup>. (Leech, Raux-Deery et al. 2002; Raux, Leech et al. 2003). The other enzymes needed to complete the synthesis of sirohaem in *B. megaterium* have also been identified; SirC, a PC-2 dehydrogenase and SirB, a ferrochelatase. Subsequently, the isofunctional *B. subtilis* enzymes are renamed: YInD as SirA, YInF as SirC and YInE as SirB (Raux, Leech et al. 2003).

### 1.3.3.4. Met1p in *Saccharomyces cerevisiae* in the sirohaem pathway

In *S. cerevisiae* two genes have been identified for the transformation of UIII into sirohaem, *MET1* and *MET 8*. Sirohaem is used exclusively in sulfite reduction, as this organism does not possess a nitrite reductase. (Hansen, Muldbjerg et al. 1997; Raux, McVeigh et al. 1999). *MET1* encodes the SUMT, Met1p (~65kDa) that has 593 amino acids, larger than the 457 amino acid CysG. Sequence comparisons showed the C-terminus (amino acids 326-556) is similar to the domain of CysG conferring SUMT activity, CysG<sup>A</sup>. *MET8* encodes for the bifunctional enzyme Met8p (~32kDa), a NAD<sup>+</sup>-dependent PC-2 dehydrogenase and a ferrochelatase. The structure of Met8p was resolved to 2.2 Å in 2002 which also portrayed a homodimeric structure (Schubert, Raux et al. 2002).

### 1.3.3.5. CorA in *Methanobacterium ivanovii*

CorA is also a homodimer with a subunit mass of 29kDa found to have SUMT activity in *M. ivanovii*. In contrast to CobA from *B. megaterium* and *Ps. denitrificans* that exhibited low [UIII] substrate inhibition, CorA from *M. ivanovii* is inhibited by [UIII] of close to 20 μM. It is perceived that since methanogens require

such high levels of tetrapyrrole coenzyme F<sub>430</sub> that inhibition does not occur at low UIII substrate concentrations, to allow greater flux through the pathway (Blanche, Robin et al. 1991).

#### **1.3.3.6. UPM1 from *A. thaliana* and ZmSUMT1 from maize**

Tetrapyrrole biosynthesis in plants is crucial to plant metabolism (Tanaka and Tanaka 2007). Higher plants such as *A. thaliana* synthesize the major tetrapyrroles chlorophyll, haem, and sirohaem. UIII is either methylated and iron is inserted to form sirohaem or it is oxidatively decarboxylated to form protoporphyrin IX. Insertion of Fe<sup>2+</sup> into protoporphyrin IX leads to haem whereas insertion of Mg<sup>2+</sup> is leads to the chlorophyll branch. (Vavilin and Vermaas 2002). Sirohaem is the cofactor for two chloroplast enzymes nitrite and sulphite reductases (Murphy, Siegel et al. 1974). All tetrapyrroles are synthesized in the chloroplast, or plastids in non-photosynthetic cells, with the exception that the last two steps of haem synthesis are also found in mitochondria (Smith and al 1999).

UPM1p from *A. thaliana* has been isolated and identified as an SUMT as part of the sirohaem biosynthetic pathway and shown to be ~40kda conferring ~43% homology to *E. coli* CysG<sup>A</sup>. A sequence has been identified to resemble a transit peptide for localization to plastids. The protein produced by *in vitro* expression is able to enter isolated intact chloroplasts but not mitochondria. Sirohydrochlorin chelatase At-SirB is also localised in the plastids. The search for a sirohydrochlorin dehydrogenase continues to complete the sirohaem pathway in *A. thaliana* (Leustek, Smith et al. 1997; Raux-Deery, Leech et al. 2005).

Similarly, ZmSUMT1 from maize was isolated. The purified mature ZmSUMT1 revealed a molecular mass of 52kDa using gel-filtration chromatography. It confers about 50 % homology to the *E. coli* CysG<sup>A</sup>. ZmSUMT1 was also shown to be transported to plastids, indicating the synthesis of sirohaem occurs in plastids too but further analysis is required to deduce how sirohaem is manufactured (Sakakibara, Takei et al. 1996; Fan, Wang et al. 2006).

Clearly, the presence of different SUMT enzymes specific to each pathway is a highly evolved mechanism of regulation of tetrapyrrole biosynthesis, and therefore flux through these pathways.

#### 1.3.4. The biosynthesis of haem $d_1$

As already mentioned, the synthesis of haem  $d_1$  proceeds from ALA via UIII as with all other tetrapyrroles. The precise mechanism of transformation from UIII to haem  $d_1$  poses a great challenge and is less well characterised. A distinct lack of information exists regarding the prerequisite intermediates for the formation of haem  $d_1$  and as to how the aforementioned atypical substituents are introduced into the macrocycle. Of particular interest is the incorporation of the oxygen onto the macrocycle under anaerobic conditions. As a direct result resolving the synthesis of haem  $d_1$  has attracted much interest, but no true intermediates have been isolated, but is very desirable.

Examination of both structures reveal the transformation from UIII to haem  $d_1$  necessitates methylation at C2 and C7; decarboxylation of the acetate side chains at C12 and C18; loss of the propionic side chains at C3 and C8 with subsequent oxidation at C3 and C8; dehydrogenation of the C17 propionate side chain to give the acrylate substituent and ferrochelation (Figure 1.23).

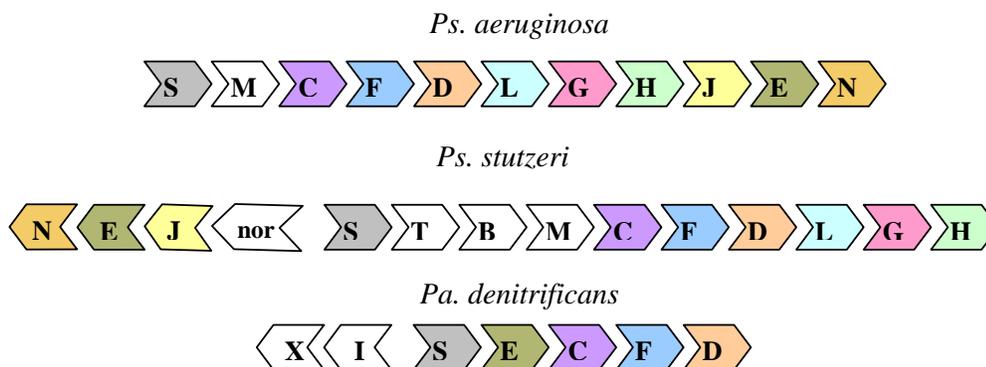
Mutation and complementation analysis of the *nir* genes have implicated a requirement of seven *nir* genes for haem  $d_1$  biogenesis that are found as an operon, called *nirE*, *nirJ*, *nirH*, *nirG*, *nirL*, *nirD* and *nirF*. Mutational analysis in *Ps. stutzeri* revealed that if any one or all *nir* genes are defective it results in an inactive semi apoform of cytochrome  $cd_1$ , lacking haem  $d_1$  but containing haem *c* (Palmedo, Seither et al. 1995). In *Ps. aeruginosa* mutational analysis as well as complementation, both implicated the same requirement of the seven *nir* genes for haem  $d_1$  biosynthesis (Kawasaki, Arai et al. 1995; Kawasaki, Arai et al. 1997).

Reconstitution of cytochrome  $cd_1$  with native haem  $d_1$  has been shown to restore nitrite reductase activity in both *Ps. stutzeri* and *Ps. aeruginosa* by 82 % and 95 % respectively. While reconstitution with synthetic haem  $d_1$  has been shown to restore nitrite reductase activity to 77 % in *Ps. stutzeri*. Significantly, no other haem can substitute for haem  $d_1$ , except for haem *a* but only 5 % activity is restored (Hill and Wharton 1978). Reconstitution with synthetic haem  $d_1$  lacking the oxo groups showed that removal of only one carbonyl resulted in complete loss of activity and removal of the acrylate resulted in a loss 60 % activity (Chang 1994). Clearly, the stereo specificity of the haem  $d_1$  macrocycle and its unique peripheral groups are relevant to the enzymes performance.

In *Ps. stutzeri* the genes for haem  $d_1$  biosynthesis are located over two loci, *nirJEN* and *nirCFDLGH*, which respectively lie upstream and downstream of the structural NiR gene *nirS* (Palmedo, Seither et al. 1995). In *Ps. aeruginosa* the same genes, *nirCFDLGHJEN* are found in one locus downstream of *nirS* and may encode the entire set of proteins required for the formation of haem  $d_1$  (Kawasaki, Arai et al. 1995; Kawasaki, Arai et al. 1997). In *Pa. denitrificans*, the presence of *nirECFD* downstream of *nirS* has been well documented (Baker, Ferguson et al. 1998; Philippot 2002). It is further proposed that NirF and NirD are fused within this organism (de Boer, Reijnders et al. 1994; Palmedo, Seither et al. 1995). The *nir* operons are diagrammatically represented in Figure 1.22.

**Figure 1.22: The organisation of haem  $d_1$  genes within the operon.**

An illustration of the *nir* operon from *Ps. aeruginosa*, *Ps. stutzeri* and *Pa. denitrificans*. Block arrows show the direction of transcription and homologues are shown in the same colour. The genes have not been drawn to scale. Nir I and X are shown to be involved in transcriptional regulation of the operon.



The discovery of the genes required in the biosynthesis of haem  $d_1$  was significant. However, in the past ten years, very little has been revealed about the structure and function of the proteins that they encode, and indeed any intermediates they may synthesise. Comparative analysis to known proteins of some similarity gives little indication as to their structure and the reactions they may catalyse. Therefore resolving the structure of these proteins would be advantageous, especially in determining function and furthermore any intermediates they may catalyse in the haem  $d_1$  pathway.

#### 1.3.4.1. NirE as a SUMT

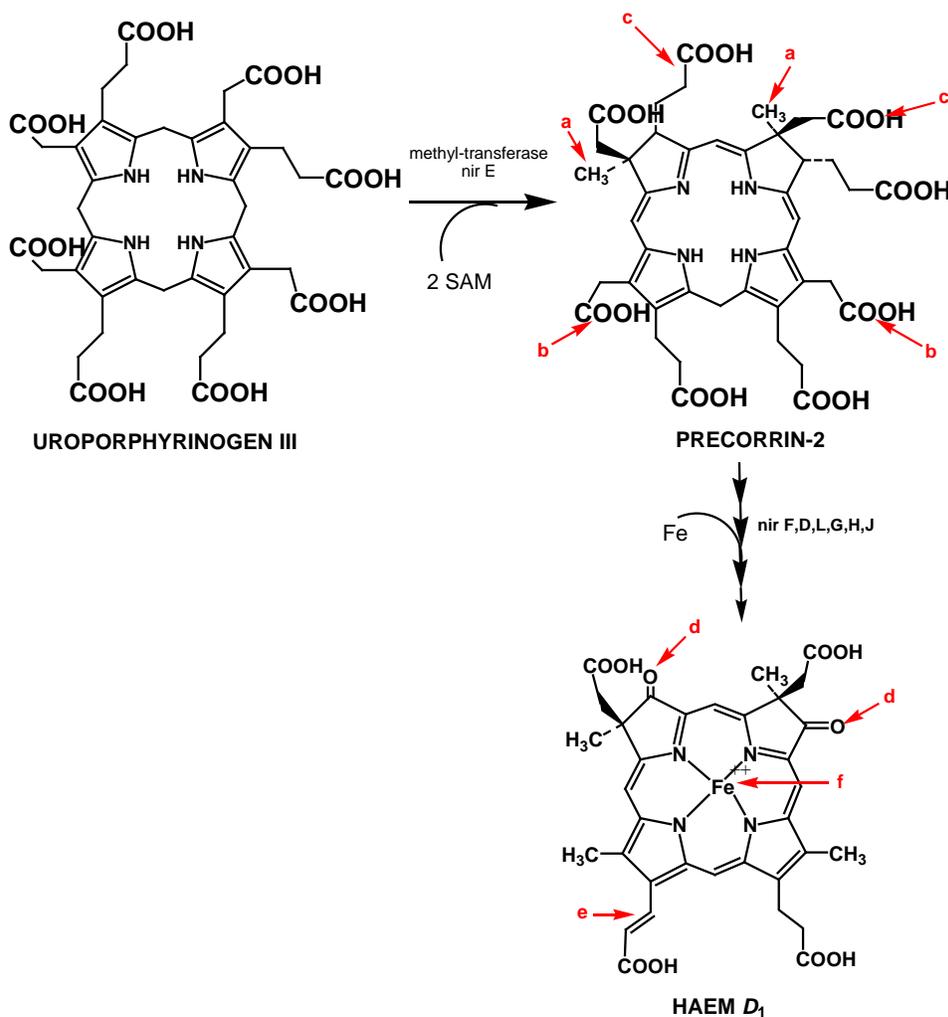
It is reported that the gene product of *nirE* encodes for a SUMT. This is supported by the similarity it shares with other methyltransferases that have been purified and characterised from different sources such as CobA and CysG (Blanche, Debussche et al. 1989; Blanche, Robin et al. 1991; Vevodova, Graham et al. 2004). *nirE* was first discovered in *Pa. denitrificans*, downstream of *nirS* (de Boer, Reijnders et al. 1994). It has since been found in *Ps. aeruginosa* (Kawasaki, Arai et al. 1997; Vollack, Xie et al. 1998) and in *Ps. stutzeri* (Palmedo, Seither et al. 1995).

Although sequence identity indicates that *nirE* encodes an S-adenosyl-L-methionine dependent uroporphyrinogen III methyltransferase, the activity of NirE has never been proven to transform UIII to PC-2. Only from mutational and complementation studies, has it been shown that *nirE* is essential for haem  $d_1$  as the lack of it perturbs biosynthesis (Palmedo, Seither et al. 1995; Kawasaki, Arai et al. 1997). The growth of *Ps. aeruginosa* on [ $^{13}\text{C}$ ] methionine, indicated methionine to be the source of the methyl groups at C2 and C7 and that C-methylation occurs as part of the route to haem  $d_1$  (Yap-Bondoc, Bondoc et al. 1990). The methylation of UIII to form PC-2 is common to the vitamin B<sub>12</sub> and sirohaem pathways suggesting UIII is the pre-cursor haem  $d_1$ , and that its synthesis is therefore likely to proceed via PC-2. However, the pathway from PC-2 remains unclear (Figure 1.23).

Experiments using labelled ALA showed the origin of the haem  $d_1$  macrocycle, including the carbonyls at C3 and C8 is derived from ALA only, except for the C2 and C7 methyl groups that are from SAM (Matthews 1993). The exact stereospecificity of the methyl groups on haem  $d_1$  was found to be exactly the same as vitamin B<sub>12</sub>, sirohaem and its intermediates, in the absolute configuration of 2*R*, 7*R*. This further substantiates a connection to the tetrapyrrole pathway implicating that methylation occurs as part of the haem  $d_1$  pathway (Micklefield 1997).

**Figure 1.23: The transformation from UIII to form haem  $d_1$** 

Analysis reveals that the transformation from UIII, is catalysed by NirE, and involves methylation at C2 and C7 (a), yielding PC-2. Further modifications include: (b) decarboxylation of the acetate side chains at C12 and C18; (c) loss of the propionic side chains at C3 and C8; (d) subsequent oxidation at C3 and C8; (e) dehydrogenation of the C17 propionate side chain to give the acrylate substituent; (f) ferrochelation.



### 1.3.4.2. The enigma of haem $d_1$ biosynthesis.

Recent papers have reported many different hypotheses on the pathway for haem  $d_1$  from UIII. Each is different and forms its own conclusions with a few similarities. One looks at a possible intermediate whilst another forms very different conclusions on the enzymes required to synthesise haem  $d_1$ .

Only one putative intermediate thus far has been isolated (Youn, Liang et al. 2004). This species, termed Compound 800, shown in Figure 5.1 (page 171), after the unique mass of its methyl ester derivative, has not been constructively put into a possible sequence of events in the pathway to haem  $d_1$ . They had initially assumed that NirFDLGH from *Ps. stutzeri* would be sufficient for haem  $d_1$  synthesis in the correct host such as *E. coli* and checked for haem  $d_1$  precursors assuming that *E. coli* CysG would be sufficient for SUMT activity, but it was to no avail. The addition of cell extract from *E. coli* containing overproduced CysG did not produce any new tetrapyrroles. Similarly, *Ps. aeruginosa* with the NirFDLGH plasmid had no effect on the level of haem  $d_1$  synthesis but a new tetrapyrrole, C800, was observed when cell extracts of *Ps. aeruginosa* with NirFDLGH and cell extract of SUMT from *D. vulgaris* were incubated anaerobically with SAM and UIII. Interestingly, there is no mention of the use of NirE as the SUMT, particularly as it has been well documented to be the preferred SUMT for this pathway and thus the use of *D. vulgaris* SUMT is debatable. NirJ has been omitted in these investigations, but it is clearly necessary for haem  $d_1$  biosynthesis. Manipulation of C800 with strong acid and heat converted it to haem  $d_1$ , however these are non-physiological conditions and will not occur otherwise. Furthermore, the structure of C800 has atypical substituents such as epoxide and sulfoxide moieties that have not previously been found in any tetrapyrrole intermediate. Therefore, the presence of C800 is somewhat new but the ambiguity of this compound makes it very difficult to confirm if it is part of, or related to a true intermediate in the pathway.

More recently, bioinformatic analysis in heliobacter (*Heliobacillus mobilis* and *Heliophilium fasciatum*) has been performed (Xiong, Bauer et al. 2007). The presence of a haem  $d_1$  like operon has been proposed, containing NirD, NirL and two forms of NirJ (NirJ1 and NirJ2) alongside genes for the production of UIII from ALA as well as CysG<sup>A</sup> and CysG<sup>B</sup>. The pathway for haem  $d_1$  has been proposed, however the acceptability of this hypothesis is debatable. Firstly not all of the

transformations from UIII to haem  $d_1$  can be explained by the presence of these enzymes. It has been hypothesised that CysG<sup>A</sup> acts as a SUMT to yield PC-2 and CysG<sup>B</sup> yields SHC, and that only NirJ is involved in decarboxylation and then CysG<sup>B</sup> is involved again in ferrochelation completing the synthesis of haem  $d_1$ . This hypothesis does not take into account the formation of an acrylate at C17 nor the conversion of the propionates to oxo groups at C3 and C8. There are no enzymes with similarity to NirG and NirH in this bacterium; this is unusual since they occur in all *nir* operons characterised thus far. Furthermore, the presence of NirD and NirL has been eliminated from the pathway and stated to be involved only in transcriptional regulation and DNA binding. This does not take into account the previously demonstrated necessity of *nirFDLGHJE* in the pathway. If the genes required for haem  $d_1$  were present as hypothesised by Xiong (NirJ1 and NirJ2; CysG<sup>A</sup> and CysG<sup>B</sup>), the insertional mutations of *nirD* or *nirL* would have no effect on haem  $d_1$  synthesis, as a lack of any one of these would be complemented by *nirG* or *nirH* (Chapter 4). It can also be said that the transcription regulation of the *nir* operon is mediated upstream of the cluster by FNR-like transcriptional activators in *Ps. aeruginosa* (ANR (Sawers 1991)) and *Pa. denitrificans* (NNR, NirI, NirX (Saunders, Houben et al. 1999)), whose transcription is mediated by nitric oxide signalling (Van Spanning, Houben et al. 1999). Nitric oxide signalling has also been implicated in transcriptional control in *Ps. aeruginosa* and *Ps. stutzeri* (Arai, Kodama et al. 1999; Vollack and Zumft 2001). Thus the transcription of these genes is tightly regulated by ANR or NNR, which in turn are activated in response to anaerobic stimuli. Therefore, the inferences made by Xiong et al can be said to have no foundation in the  $d_1$  pathway as they are based on bioinformatic analysis and manipulation only. They are not experimentally derived. The pathway proposed in this case is more likely to be related to a primitive haem pathway observed in *D. vulgaris* and *Methanosarcina barkeri*, whereby UIII is methylated instead of decarboxylated (Ishida, Yu et al. 1998; Buchenau, Kahnt et al. 2006). In *M. barkeri* the late haem genes are missing, and it is presumed that an alternative pathway is utilised to make haem. Thus, it is likely to be the same in heliobacteria. Moreover, heliobacter contain bacteriochlorophyll and have no indication of a requiring an alternative respiratory nitrogen metabolism pathway such as one requiring haem  $d_1$ .

Therefore, the pathway to haem  $d_1$  still remains to be established and the isolation of any intermediates is highly sought-after.

#### 1.4. *Ps. aeruginosa* as a candidate species

*Ps. aeruginosa* is a gram-negative bacterium found in a large variety of habitats that include soil, marine as well as plant and animal tissue. Its ubiquitous nature has allowed it to evolve to possess a large spectrum of biochemical pathways and the metabolic versatility of which *pseudomonads* are renowned. *Ps. aeruginosa* is the most thoroughly characterised of the denitrifying bacteria as a result of the interest in it as an opportunistic pathogen, as well as a versatile soil bacterium. Its ability to utilise and adapt regulatory systems has enabled it to thrive in many extreme environments, conferring resistance to many antibiotics and disinfectants. The genome has been fully sequenced revealing a large bacterial genome, 6.3 million base pairs, in comparison to other microorganisms that are near to 3 million base pairs. Clearly, its evolution as a complex bacterium is reflected in its diversity as well as genome size (Stover, Pham et al. 2000).

Therefore, the reasons for studying haem  $d_1$  biosynthesis in *Ps. aeruginosa* are multiple, including its large size, ease of genetic manipulation and its inherent neat genetic organisation. In particular, the genes *nirCFDLGHJEN* are found in one locus being read in the same direction as oppose to *Ps. stutzeri* (Figure 1.22). These may encode the entire set of proteins required for the formation of haem  $d_1$ . It is documented to have a basal level of haem  $d_1$  in wild type *Ps. aeruginosa* at 0.7 nmol per gram wet weight (Youn, Liang et al. 2004). Indeed, the use of *Ps. aeruginosa* in this study is as a model organism for understanding not only how haem  $d_1$  is synthesised, but also how it is regulated and controlled in relation to other tetrapyrroles. Such studies could lead to the conception of new tetrapyrrole intermediates that may involve new and unique chemistry in the world of tetrapyrroles.

### 1.5. Purpose of thesis

From the review of literature, we can see a clear scarceness in the information available on the pathway from UIII to haem  $d_1$ . There are no true characterised intermediates to date for the last 25 years, and only a methyl ester, C800 has been put forward. The genes necessary for the synthesis of haem  $d_1$  have been found, but little is known of the proteins they encode and their structures or mechanism of action.

Thus, the purpose of the research conducted was to elucidate the pathway for the production of haem  $d_1$ , including the characterisation of the enzymes and any intermediates. This thesis hoped to resolve some of the questions and challenges posed in this pathway.

In this thesis, a number of different approaches are employed in the study of such a unique pathway. In Chapter 3, it was confirmed for the first time that NirE was a true SUMT and shown to be active by catalysing the transformation from UIII to PC-2, with the formulation of an anaerobic assay coupled to the formation of sirohydrochlorin. Chapter 4 shows the concerted effort required in the optimisation of overproduction and stabilisation of the remaining Nir proteins and the conception of a novel homologous system of expression in *Ps. aeruginosa*. The last chapter focuses on engineering both *E. coli* and *Ps. aeruginosa* to make haem  $d_1$  and thus, the expression of intermediates by use of Link and Lock cloning, to clone the *nir* genes consecutively into a plasmid. The accumulation of intermediates was also examined by analysis of activity of the Nir proteins towards PC-2 and SHC.

Therefore, this thesis provides a suitable foundation for future work in the investigation of the haem  $d_1$  tetrapyrrole pathway, its enzymes and intermediates.

## CHAPTER 2

---

# **Materials and Methods**

## 2. Materials and Methods

### 2.1. Materials

#### 2.1.1. Chemicals (source)

Almost all chemicals such as antibiotics and tetrapyrroles were purchased from Sigma-Aldrich Ltd (Poole, UK) unless otherwise stated. Restriction enzymes were purchased from Promega (Madison, USA), except for *AscI* that was purchased from New England Biolabs (Hitchin, UK); pGEMT®-Easy vector systems and the bacterial strain JM109 were also purchased from Promega (Madison, USA); pET14b and pETcoco cloning vectors were purchased from Novagen; Bacterial Agar, Tryptone and yeast extract were purchased from Oxiod Ltd (Basingstoke, UK); Hot star Taq DNA polymerase, MiniElute™ Gel extraction Kit, QIAprep® Spin Miniprep Kit and QIAquick® Gel Extraction Kit from Qiagen GmbH (Germany); PD10 columns, Sephadex G-75, Sepharose S-200, DEAE Sephacel and chelating sepharose fast flow resin were from Amersham Biosciences (Amersham Pharmacia, UK); 50 x TAE Buffer was purchased from Eppendorf (Cambridge, UK) and IPTG from BDH (VWR international Ltd, Poole, UK).

#### 2.1.2. Bacterial Strains

Bacterial Strains JM109, BL21 (subtypes) were purchased from Novagen (UK), Promega (USA) and SCS110 was provided by Stratagene (USA), except for those carrying the prefix ER. Those strains carrying the prefix ER were provided courtesy of Dr Evelyne Raux-Deery (University of Kent, Canterbury). The list of cloning strains used in this study can be found in Table 2.1. All other strains made with the desired plasmid in *Ps. aeruginosa* or in *E. coli* are listed in Table 2.2.

#### 2.1.3. Plasmids

The vectors used in this study have been listed in Table 2.3 and the plasmids that were constructed as part of the research are listed in Table 2.4. The plasmids with a prefix pER were given by Dr Evelyne Raux-Deery (University of Kent, Canterbury).

**Table 2.1: List of stock strains used in this study**

All genes in the bacterium are presumed to be in the wild type state, except for those listed, which are mutant alleles carried by that bacterium. Genes listed on the F' episome represent wild type alleles unless specified otherwise. Strains are  $\lambda^-$  unless specified otherwise. A description of genotype or phenotype incurred is detailed in the appendix.

Strains	Genotype / Phenotype	Description	Reference/ Source
JM109	<i>endA1 recA1 gyrA96, thi hsdR17(r<sub>K</sub>-m<sub>K</sub><sup>+</sup>) relA1 supE44 Δ(lac-proAB) [F', traD36 proAB lacI<sup>q</sup>ZΔM15]</i>	Cloning strain used for simple molecular biology applications	Promega / (Hanahan 1983; Hanahan 1985)
BL21(DE3) pLysS	F <sup>-</sup> , <i>ompT</i> , <i>hsdS<sub>B</sub></i> (r <sub>B</sub> <sup>-</sup> , m <sub>B</sub> <sup>-</sup> ), <i>dcm</i> , <i>gal</i> , λ(DE3), pLysS, Cm <sup>R</sup>	Protein expression strain for plasmids utilising the T7 promoter system. Contains λ DE3 lysogen carrying gene for T7 RNA polymerase under control of the lacUV5 promoter. IPTG induces expression of the T7 RNA polymerase.	Novagen
SCS110	<i>RpsL</i> (Str <sup>r</sup> ) <i>thr leu endA thi-l galK galT ara tonA tsx dam dcm supE44 Δ(lac-proAB) [F', traD36 lacI<sup>q</sup>ZΔM15] proAB, Cm<sup>R</sup></i>	An <i>endA</i> <sup>-</sup> derivative of the JM110 strain. Used for preparing plasmid or phagemid DNA, free of Dam or Dcm methylation allowing restriction by one or more methylation-sensitive restriction enzymes.	Stratagene
BL21 Star (DE3)	F <sup>-</sup> <i>ompT hsdSB</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>gal dcm rne131</i> (DE3) (Cm <sup>R</sup> )	Contains a mutation in the gene encoding RNaseE ( <i>rne131</i> ), one of the major sources of mRNA degradation. Improves stability of mRNA transcripts and increases protein expression yield from T7 promoter-based vectors. Do not contain the <i>lon</i> protease or the outer membrane protease, <i>OmpT</i> , further reducing degradation of heterologous proteins expressed in the strains	Invitrogen
BL21star (DE3)pLysS	F <sup>-</sup> <i>ompT hsdSB</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>gal dcm rne131</i> (DE3) pLysS (Cm <sup>R</sup> )	Same as BL21 Star (DE3) but BL21 Star (DE3) pLysS carries the pLysS plasmid which produces T7 lysozyme	Invitrogen
AR3612 (pCIQ)	<i>S. enterica</i> ( <i>typhimurium</i> ) Leu <sup>+</sup> Sm <sup>R</sup> <i>cysG</i> <sup>A-</sup> MetE, Cm <sup>r</sup>	Lacks CysG <sup>A</sup> activity, but retains CysG <sup>B</sup> activity. <i>Met H</i> B <sub>12</sub> dependent and B <sub>12</sub> deficient	(Raux, Lanois et al. 1996)

TUNER	$F^- ompT hsdS_B (r_B^- m_B^-)$ <i>gal dcm lacY1</i> (DE3) pLacI (Cm <sup>R</sup> )	<i>lacZY</i> deletion mutants of BL21, enabling adjustable levels of protein expression throughout all cells in a culture. The lac permease ( <i>lacY</i> ) mutation allows uniform entry of IPTG into all cells in population, producing a concentration-dependent, homogeneous level of induction. By adjusting the concentration of IPTG, expression can be finely regulated in pET vectors	Invitrogen
ER171	<i>cysG</i> 302Δa (pAR8068) pACYC184 <i>lacI</i> <sup>q</sup>	<i>E. coli cysG</i> deleted strain	Dr E.Raux-Deery
CysG- 247	JM101 pCIQ pTac <i>cobA</i>	<i>E. coli cysG</i> deleted strain, with <i>cobA</i> from <i>Ps. denitrificans</i>	Dr E.Raux-Deery
<i>Pseudomonas aeruginosa</i>	PAO1	<i>Ps. aeruginosa</i> wt	(Dunn and Holloway 1971)

**Table 2.2: List of strains used as part of this study**

A library of strains containing the various single and multigene plasmids was compiled using strains from both *Ps. aeruginosa* (PAO wt) and *E. coli* and stored as glycerol stocks at  $-80\text{ }^{\circ}\text{C}$  (see Table 2.4 for plasmids).

Strains	Genotype / Phenotype	Description	Reference/ Source
VP01	BL21(DE3) pLysS (pVP14)	<i>Ps. aeruginosa</i> NirF Histagged	This study
VP02	BL21(DE3) pLysS (pVP13)	<i>Ps. aeruginosa</i> NirD Histagged	This study
VP03	BL21(DE3) pLysS (pVP09)	<i>Ps. aeruginosa</i> NirL Histagged	This study
VP04	BL21(DE3) pLysS (pVP05)	<i>Ps. aeruginosa</i> NirG Histagged	This study
VP05	BL21(DE3) pLysS (pVP06)	<i>Ps. aeruginosa</i> NirH Histagged	This study
VP06	BL21(DE3) pLysS (pVP66)	<i>Ps. aeruginosa</i> NirJ Histagged	This study
VP07	BL21(DE3) pLysS (pVP02)	<i>Ps. aeruginosa</i> NirE Histagged	This study
VP08	BL21(DE3) pLysS (pVP71)	<i>Ps. aeruginosa</i> NirDL (NirD Histagged)	This study
VP09	BL21(DE3) pLysS (pVP72)	<i>Ps. aeruginosa</i> NirDLG (NirD Histagged)	This study
VP10	BL21(DE3) pLysS (pVP73)	<i>Ps. aeruginosa</i> NirDLGH (NirD Histagged)	This study
VP11	TUNER (pVP14)	<i>Ps. aeruginosa</i> NirF Histagged	This study
VP12	TUNER (pVP13)	<i>Ps. aeruginosa</i> NirD Histagged	This study

VP13	TUNER (pVP09)	<i>Ps. aeruginosa</i> NirL Histagged	This study
VP14	TUNER (pVP05)	<i>Ps. aeruginosa</i> NirG Histagged	This study
VP15	TUNER (pVP06)	<i>Ps. aeruginosa</i> NirH Histagged	This study
VP16	TUNER (pVP66)	<i>Ps. aeruginosa</i> NirJ Histagged	This study
VP17	TUNER (pVP02)	<i>Ps. aeruginosa</i> NirE Histagged	This study
VP18	PAO1 (pVP19)	<i>Ps. aeruginosa</i> NirF Histagged	This study
VP19	PAO1 (pVP21)	<i>Ps. aeruginosa</i> NirD Histagged	This study
VP20	PAO1 (pVP47)	<i>Ps. aeruginosa</i> NirL Histagged	This study
VP21	PAO1 (pVP23)	<i>Ps. aeruginosa</i> NirG Histagged	This study
VP22	PAO1 (pVP48)	<i>Ps. aeruginosa</i> NirH Histagged	This study
VP23	PAO1 (pVP78)	<i>Ps. aeruginosa</i> NirJ Histagged	This study
VP24	PAO1 (pVP17)	<i>Ps. aeruginosa</i> NirE Histagged	This study
VP25	PAO1 (pVP18)	<i>Ps. aeruginosa</i> NirF	This study
VP26	PAO1 (pVP20)	<i>Ps. aeruginosa</i> NirD	This study
VP27	PAO1 (pVP25)	<i>Ps. aeruginosa</i> NirL	This study
VP28	PAO1 (pVP22)	<i>Ps. aeruginosa</i> NirG	This study
VP29	PAO1 (pVP24)	<i>Ps. aeruginosa</i> NirH	This study
VP30	PAO1 (pVP77)	<i>Ps. aeruginosa</i> NirJ	This study
VP31	PAO1 (pVP16)	<i>Ps. aeruginosa</i> NirE	This study
VP32	PAO1 (pVP33)	<i>Ps. aeruginosa</i> NirE (Nde/Xba/BamHI primers)	This study
VP33	PAO1 (pVP50)	<i>Ps. aeruginosa</i> NirEF	This study
VP34	PAO1 (pVP51)	<i>Ps. aeruginosa</i> NirEFD	This study
VP35	PAO1 (pVP52)	<i>Ps. aeruginosa</i> NirEFDL	This study
VP36	PAO1 (pVP53)	<i>Ps. aeruginosa</i> NirEFDLG	This study
VP37	PAO1 (pVP54)	<i>Ps. aeruginosa</i> NirEFDLGH	This study
VP38	PAO1 (pVP55)	<i>Ps. aeruginosa</i> NirEFDLGHJ	This study
VP39	PAO1 (pVP58)	<i>Ps. aeruginosa</i> NirF	This study
VP40	PAO1 (pVP59)	<i>Ps. aeruginosa</i> NirFD	This study
VP41	PAO1 (pVP60)	<i>Ps. aeruginosa</i> NirFDL	This study
VP42	PAO1 (pVP61)	<i>Ps. aeruginosa</i> NirFDLG	This study
VP43	PAO1 (pVP62)	<i>Ps. aeruginosa</i> NirFDLGH	This study
VP44	PAO1 (pVP63)	<i>Ps. aeruginosa</i> NirFDLGHJ	This study
VP45	PAO1 (pVP65)	<i>Ps. aeruginosa</i> NirFDLGHJE	This study

VP46	<i>E. coli</i> JM109 (pVP33)	<i>Ps. aeruginosa</i> NirE	This study
VP47	<i>E. coli</i> JM109 (pVP50)	<i>Ps. aeruginosa</i> NirEF	This study
VP48	<i>E. coli</i> JM109 (pVP51)	<i>Ps. aeruginosa</i> NirEFD	This study
VP49	<i>E. coli</i> JM109 (pVP52)	<i>Ps. aeruginosa</i> NirEFDL	This study
VP50	<i>E. coli</i> JM109 (pVP53)	<i>Ps. aeruginosa</i> NirEFDLG	This study
VP51	<i>E. coli</i> JM109 (pVP54)	<i>Ps. aeruginosa</i> NirEFDLGH	This study
VP52	<i>E. coli</i> JM109 (pVP55)	<i>Ps. aeruginosa</i> NirEFDLGHJ	This study
VP53	PAO1 (pVP67)	<i>Ps. aeruginosa</i> NirDL	This study
VP54	PAO1 (pVP68)	<i>Ps. aeruginosa</i> NirDLG	This study
VP55	PAO1 (pVP69)	<i>Ps. aeruginosa</i> NirDLGH	This study
VP56	PAO1 (pVP71)	<i>Ps. aeruginosa</i> NirDL (NirD Histagged)	This study
VP57	PAO1 (pVP72)	<i>Ps. aeruginosa</i> NirDLG (NirD Histagged)	This study
VP58	PAO1 (pVP73)	<i>Ps. aeruginosa</i> NirDLGH (NirD Histagged)	This study

**Table 2.3: List of stock vectors used in the study**

Cloning vectors and plasmids obtained from laboratory members and companies.

Vector	Description	Source
pGEM-T Easy	T-vector cloning system for PCR products; MCS contains 3'-T overhangs for ease of ligation. Amp <sup>R</sup> , T7, SP6 RNA polymerase promoters	Promega
pET14b	Protein overproduction vector with N-terminal His-tag. Amp <sup>R</sup> , T7 promoter, T7 terminator	Novagen
pUCP-Nco	NcoI (C/CATGG) in MCS downstream of promoter. Amp <sup>R</sup> , <i>lac</i> promoter	Warren Lab (Ross Graham) (Cronin and McIntire 1999)
pUCP-Nde	NdeI (CA/TATG) in MCS downstream of promoter. Amp <sup>R</sup> , <i>lac</i> promoter	Warren Lab (Ross Graham)(Cronin and McIntire 1999)
pKK233.3 + <i>haem d<sub>1</sub></i> operon	<i>nirF</i> , <i>nirD</i> , <i>nirL</i> , <i>nirG</i> , <i>nirH</i> , <i>nirJ</i> , <i>nirE</i> cloned as part of operon. Amp <sup>R</sup>	Donated by Evelyne Deery
<i>nirJ</i> pET14b	<i>nirJ</i> with N-terminal His-tag, Amp <sup>R</sup> , T7 promoter, T7 terminator	Donated by Andrew Lawrence (mutation)
<i>cobA</i> pET14b	<i>cobA</i> with N-terminal His-tag, Amp <sup>R</sup> , T7 promoter, T7 terminator	Donated by Amanda Brindley
pETcoco ABCD	<i>hemB</i> , <i>hemC</i> , <i>hemD</i> , <i>cobA</i> with N-terminal His-tag. Amp <sup>R</sup> , T7 promoter, T7 terminator (for PC-2)	Donated by Evelyne Deery
pETcoco ABCDC	<i>hemB</i> , <i>hemC</i> , <i>hemD</i> , <i>cobA</i> , <i>sirC</i> with N-terminal His-tag, Amp <sup>R</sup> , T7 promoter, T7 terminator (for UIII)	Donated by Evelyne Deery
pET14b BCDC	<i>hemB</i> , <i>hemC</i> , <i>hemD</i> , <i>sirC</i> with N-terminal His-tag. Amp <sup>R</sup> , T7 promoter, T7 terminator (for SHC no SUMT)	Donated by Amanda Brindley

**Table 2.4: List of plasmids constructed and used in this study**

Unless otherwise stated, all the genes were cloned from *Pseudomonas aeruginosa*, RBS refers to ribosome binding site. A total of 78 plasmids were contrasted as part of this study.

Construct	Gene	Vector	Description / Origin	Mutation free
<b>Single gene construct: <i>NdeI</i> – <i>BamHI</i> Primers in pGEMT-Easy and pET14b</b>				
pVP11	<i>nirF</i>	pGEMT-E	PCR1 fragment to introduce <i>NdeI</i> and <i>BamHI</i> primer sites (clones with two gene orientations in plasmid)	<b>D1039N</b>
pVP12	<i>nirF</i>	pET14b	<i>NdeI</i> - <i>BamHI</i> fragment from pVP11 PCR1	<b>D1039N</b>
pVP14	<i>nirF</i>	pET14b	PCR2 fragment <i>NdeI</i> and <i>BamHI</i> primer sites, direct cloning to vector	<b>Yes</b>
pVP07	<i>nirD</i>	pGEMT-E	PCR1 fragment to introduce <i>NdeI</i> and <i>BamHI</i> primer sites	<b><i>NdeI</i> lost</b>
pVP10	<i>nirD</i>	pET14b	<i>NdeI</i> - <i>BamHI</i> fragment from pVP07 PCR1	<b><i>NdeI</i> lost</b>
pVP13	<i>nirD</i>	pET14b	PCR2 fragment <i>NdeI</i> and <i>BamHI</i> primer sites, direct cloning to vector	<b>Yes</b>
pVP08	<i>nirL</i>	pGEMT-E	PCR1 fragment to introduce <i>NdeI</i> and <i>BamHI</i> primer sites	<b>Yes</b>
pVP09	<i>nirL</i>	pET14b	<i>NdeI</i> - <i>BamHI</i> fragment from pVP08 PCR1	<b>Yes</b>
pVP03	<i>nirG</i>	pGEMT-E	PCR1 fragment to introduce <i>NdeI</i> and <i>BamHI</i> primer sites	<b>Yes</b>
pVP05	<i>nirG</i>	pET14b	<i>NdeI</i> - <i>BamHI</i> fragment from pVP03 PCR1	<b>Yes</b>
pVP04	<i>nirH</i>	pGEMT-E	PCR1 fragment to introduce <i>NdeI</i> and <i>BamHI</i> primer sites	<b>Yes</b>
pVP06	<i>nirH</i>	pET14b	<i>NdeI</i> - <i>BamHI</i> fragment from pVP04 PCR1	<b>Yes</b>
pVP01	<i>nirE</i>	pGEMT-E	PCR1 fragment to introduce <i>NdeI</i> and <i>BamHI</i> primer sites (clones with two gene orientations in plasmid)	<b>Yes</b>

pVP02	<i>nirE</i>	pET14b	<i>NdeI-BamHI</i> fragment from pVP01 PCR1	Yes
pVP15	<i>nirJ</i>	pET14b	<i>NdeI-BamHI</i> primer sites courtesy of A Lawrence	No
pVP66	<i>nirJ</i>	pET14b	<i>AscI</i> and <i>XhoI</i> fragment from pKK Haem <i>d</i> <sub>1</sub> plasmid, cloned into pVP15	Yes
<b>Construct</b>	<b>Gene</b>	<b>Vector</b>	<b>Description / Origin</b>	<b>Mutation free</b>
<b>Single gene construct: <i>NdeI</i> or <i>NcoI</i> – <i>BamHI</i> Primers in pUCPNde or pUCPNco</b>				
pVP18	<i>nirF</i>	pUCP-Nde	<i>NdeI-BamHI</i> fragment from pVP14	Yes
pVP19	<i>nirF</i>	pUCP-Nco	<i>NcoI-BamHI</i> fragment from pVP14 + histag	Yes
pVP20	<i>nirD</i>	pUCP-Nde	<i>NdeI-BamHI</i> fragment from pVP13	Yes
pVP21	<i>nirD</i>	pUCP-Nco	<i>NcoI-BamHI</i> fragment from pVP13 + histag	Yes
pVP25	<i>nirL</i>	pUCP-Nde	<i>NdeI-BamHI</i> fragment from pVP09	Yes
pVP47	<i>nirL</i>	pUCP-Nco	<i>NdeI-BamHI</i> fragment from pVP09 + Vector and histag from pVP19 <i>NdeI-BamHI</i> prerestricted.	Yes
pVP22	<i>nirG</i>	pUCP-Nde	<i>NdeI-BamHI</i> fragment from pVP05	Yes
pVP23	<i>nirG</i>	pUCP-Nco	<i>NcoI-BamHI</i> fragment from pVP05 + histag	Yes
pVP24	<i>nirH</i>	pUCP-Nde	<i>NdeI-BamHI</i> fragment from pVP06	Yes
pVP48	<i>nirH</i>	pUCP-Nco	<i>NdeI-BamHI</i> fragment from pVP06 + Vector and histag from pVP19 <i>NdeI-BamHI</i> prerestricted.	Yes
pVP77	<i>nirJ</i>	pUCP-Nde	<i>NdeI-BamHI</i> fragment from pVP66	Yes
PVP78	<i>nirJ</i>	pUCP-Nco	<i>NcoI-BamHI</i> fragment from pVP66 + histag	Yes
pVP16	<i>nirE</i>	pUCP-Nde	<i>NdeI-BamHI</i> fragment from pVP02	Yes
pVP17	<i>nirE</i>	pUCP-Nco	<i>NcoI-BamHI</i> fragment from pVP02 + histag	Yes

Construct	Gene	Vector	Description / Origin	Mutation free
<b>Multi-gene constructs (<i>NdeI</i> or <i>NheI</i> – RBS and <i>XbaI</i> - <i>BamHI</i> Primers)</b>				
pVP27	<i>nirF</i>	pGEMT-E	PCR1 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end (2 orientations of gene)	Yes
pVP28	<i>nirD</i>	pGEMT-E	PCR1 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	G1713A, <i>NheI</i> Lost
pVP34	<i>nirD</i>	pGEMT-E	PCR2 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end (reverse orientation to PCR1)	No stop codon
pVP40	<i>nirD</i>	pGEMT-E	PCR3 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	N93S
pVP41	<i>nirD</i>	pUCPNde	PCR1 <i>NdeI</i> - <i>XbaI</i> - <i>BamHI</i> fragment from pVP28	G1713A, <i>NheI</i> Lost
pVP42	<i>nirD</i>	pUCPNco	PCR2 <i>NcoI</i> - <i>XbaI</i> - <i>BamHI</i> fragment pVP34	No stop codon
pVP43	<i>nirD</i>	pUCPNde	pVP41 vector prerestricted with <i>AvaI</i> and <i>BamHI</i> + <i>AvaI</i> and <i>BamHI</i> fragment from pVP42	Yes
pVP29	<i>nirL</i>	pGEMT-E	PCR1 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	Yes
pVP30	<i>nirG</i>	pGEMT-E	PCR1 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	H14Y, V115A
pVP35	<i>nirG</i>	pGEMT-E	PCR2 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	Very poor reading
PVP38	<i>nirG</i>	pGEMT-E	PCR3 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	A76V
PVP39	<i>nirG</i>	pGEMT-E	PCR4 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	Yes
pVP31	<i>nirH</i>	pGEMT-E	PCR1 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	Yes
pVP32	<i>nirJ</i>	pGEMT-E	PCR1 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	R36H, Q119Q

Construct	Gene	Vector	Description / Origin	Mutation free
pVP36	<i>nirJ</i>	pGEMT-E	PCR2 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	I62V
pVP37	<i>nirJ</i>	pGEMT-E	<i>AscI</i> and <i>XhoI</i> fragment from pVP15, cloned into pVP32	I62V (found in pVP15)
pVP44	<i>nirJ</i>	pGEMT-E	PCR3 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	I4T
pVP45	<i>nirJ</i>	pGEMT-E	PCR4 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	Yes
pVP46	<i>nirJ</i>	pGEMT-E	<i>AscI</i> and <i>XhoI</i> fragment from pKK Haem <i>d</i> <sub>1</sub> plasmid, cloned into pVP32	Yes
pVP26	<i>nirE</i>	pGEMT-E	PCR1 to RBS- <i>NdeI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	D29D, <i>NdeI</i> lost
pVP33	<i>nirE</i>	pGEMT-E	PCR2 to RBS- <i>NdeI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	Yes
PVP64	<i>nirE</i>	pGEMT-E	PCR1 to RBS- <i>NheI</i> sites at the 5' end and <i>XbaI</i> - <i>BamHI</i> 3' end	Yes
pVP49	<i>nirE</i>	pUCP-Nde	<i>NdeI</i> - <i>BamHI</i> fragment from pVP33	Yes
pVP50	<i>nirEF</i>	pUCP-Nde	<i>NheI</i> - <i>BamHI</i> fragment from pVP27, cloned into <i>XbaI</i> - <i>BamHI</i> of pVP49	Yes
pVP51	<i>nirEFD</i>	pUCP-Nde	<i>NheI</i> - <i>BamHI</i> fragment from pVP43, cloned into <i>XbaI</i> - <i>BamHI</i> of pVP50	Yes
pVP52	<i>nirEFDL</i>	pUCP-Nde	<i>NheI</i> - <i>BamHI</i> fragment from pVP29, cloned into <i>XbaI</i> - <i>BamHI</i> of pVP51	Yes
pVP53	<i>nirEFDL</i> <i>G</i>	pUCP-Nde	<i>NheI</i> - <i>BamHI</i> fragment from pVP39, cloned into <i>XbaI</i> - <i>BamHI</i> of pVP52	Yes
pVP54	<i>nirEFDL</i> <i>GH</i>	pUCP-Nde	<i>NheI</i> - <i>BamHI</i> fragment from pVP31, cloned into <i>XbaI</i> - <i>BamHI</i> of pVP53	Yes
pVP55	<i>nirEFDL</i> <i>GHJ</i>	pUCP-Nde	<i>NheI</i> - <i>BamHI</i> fragment from pVP45, cloned into <i>XbaI</i> - <i>BamHI</i> of pVP54	Yes

<b>Construct</b>	<b>Gene</b>	<b>Vector</b>	<b>Description / Origin</b>	<b>Mutation free</b>
pVP58	<i>nirF</i>	pUCP-Nde	<i>SalI</i> - <i>Bam</i> HI fragment from pVP27, cloned into <i>SalI</i> - <i>Bam</i> HI of pVP18	Yes
pVP59	<i>nirFD</i>	pUCP-Nde	<i>NheI</i> - <i>Bam</i> HI fragment from pVP43, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP58	Yes
pVP60	<i>nirFDL</i>	pUCP-Nde	<i>NheI</i> - <i>Bam</i> HI fragment from pVP29, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP59	Yes
pVP61	<i>nirFDLG</i>	pUCP-Nde	<i>NheI</i> - <i>Bam</i> HI fragment from pVP39, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP60	Yes
pVP62	<i>nirFDLG<sub>H</sub></i>	pUCP-Nde	<i>NheI</i> - <i>Bam</i> HI fragment from pVP31, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP61	Yes
pVP63	<i>nirFDLG<sub>HJ</sub></i>	pUCP-Nde	<i>NheI</i> - <i>Bam</i> HI fragment from pVP45, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP62	Yes
pVP65	<i>nirFDLG<sub>HJE</sub></i>	pUCP-Nde	<i>NheI</i> - <i>Bam</i> HI fragment from pVP33, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP63	Yes
pVP67	<i>nirDL</i>	pUCP-Nde	<i>NheI</i> - <i>Bam</i> HI fragment from pVP29, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP43	Yes
pVP68	<i>nirDLG</i>	pUCP-Nde	<i>NheI</i> - <i>Bam</i> HI fragment from pVP39, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP67	Yes
pVP69	<i>nirDLGH</i>	pUCP-Nde	<i>NheI</i> - <i>Bam</i> HI fragment from pVP31, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP68	Yes
pVP70	<i>nirD</i>	pUCP-Nco (+ histag)	<i>AvaI</i> - <i>Bam</i> HI fragment from pVP29, cloned into <i>AvaI</i> - <i>Bam</i> HI pVP21	Yes
pVP71	<i>nirDL</i>	pUCP-Nco (+ histag)	<i>NheI</i> - <i>Bam</i> HI fragment from pVP29, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP70	Yes
pVP72	<i>nirDLG</i>	pUCP-Nco (+ histag)	<i>NheI</i> - <i>Bam</i> HI fragment from pVP29, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP71	Yes
pVP73	<i>nirDLGH</i>	pUCP-Nco (+ histag)	<i>NheI</i> - <i>Bam</i> HI fragment from pVP29, cloned into <i>XbaI</i> - <i>Bam</i> HI of pVP72	Yes

<b>Construct</b>	<b>Gene</b>	<b>Vector</b>	<b>Description / Origin</b>	<b>Mutation free</b>
pVP74	<i>nirDL</i>	pET14b	<i>NdeI</i> - <i>BamHI</i> fragment from pVP71	Yes
pVP75	<i>nirDLG</i>	pET14b	<i>NdeI</i> - <i>BamHI</i> fragment from pVP72	Yes
pVP76	<i>nirDLGH</i>	pET14b	<i>NdeI</i> - <i>BamHI</i> fragment from pVP73	Yes
pVP56	<i>nirED</i>	pUCP-Nde	<i>NheI</i> - <i>BamHI</i> fragment from pVP43, cloned into <i>XbaI</i> - <i>BamHI</i> of pVP49	Yes
PVP57	<i>nirEG</i>	pUCP-Nde	<i>NheI</i> - <i>BamHI</i> fragment from pVP39, cloned into <i>XbaI</i> - <i>BamHI</i> of pVP49	Yes

#### 2.1.4. Media and Solutions for bacterial work

##### Luria Bertani (LB) Broth:

NaCl	5 g
Yeast	5 g
Tryptone	10 g

This was made up to 1L with distilled water and autoclaved before use.

##### Super LB Broth:

NaCl	5 g
Yeast	20 g
Tryptone	32 g

This was made up to 1L with distilled water and autoclaved before use.

##### LB Agar:

Bacterio-Agar	15 g
---------------	------

15 g of agar was added per litre of LB Broth and autoclaved before use.

##### 2TY Broth:

Peptone	16 g
Yeast	10 g
NaCl	5 g

This was made up to 1L with distilled water and autoclaved before use.

##### Minimal Media:

The solutions made up as below, were all sterilised separately and mixed together to make 1L of minimal media or broth with or without the addition of 15 g bacto-agar.

<b>Stock</b>	<b>Final Concentration</b>	<b>Volume</b>
x 10 M9 Salts	x 1 M9 Salts	100 ml
20 % Glucose	0.02 %	20 ml
1M MgSO <sub>4</sub>		2 ml
0.1M CaCl <sub>2</sub>		1 ml
Cysteine (5 mg/ml)		10 ml
H <sub>2</sub> O (+/- 15 g bacto-agar)		900 ml
+ Antibiotics as required		



20 % Glucose

Glucose (C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>) 20 % w / v filter sterilised

Dissolved in ddH<sub>2</sub>O, filter sterilised and stored at -20 °C.

IPTG Solution (1M):

IPTG 2.38 g

Dissolved in 10 mls ddH<sub>2</sub>O and filter sterilised and used in culture at a final concentration of 0.4 mM.

L-Methionine (10 mg/ml)

C<sub>5</sub>H<sub>11</sub>NO<sub>2</sub>S 0.1 g

Dissolved in 10 mls ddH<sub>2</sub>O and filter sterilised; used at a final concentration of 10 mg/l in liquid cultures.

δ-Aminoleavulinic Acid (10 mg/ml)

C<sub>5</sub>H<sub>9</sub>NO<sub>3</sub>.HCl 0.1 g

Dissolved in 10 mls ddH<sub>2</sub>O and filter sterilised; used at a final concentration of 10 mg/l in liquid cultures.

L-Cysteine (5 mg/ml)

Dissolved in 10 mls ddH<sub>2</sub>O and filter sterilised; used at a final concentration of 10 mg in complementation in 1L of media.

Sodium Nitrate (1M)

NaNO<sub>3</sub> 84.99 g

All added to 1L of dH<sub>2</sub>O and autoclaved before use in minimal media preparation.

**2.1.5. Solutions for DNA work**

All solutions were made using Analar (ddH<sub>2</sub>O)

<u>TE Buffer</u>	Tris-HCl pH 8.0	10 mM
	EDTA pH 8.0	1 mM

<u>6 x DNA Loading Buffer</u>	Bromophenol blue	0.25 %
	Glycerol	50 %
	TE Buffer	50 %

<u>TAE Buffer</u>	Tris-Acetate	40 mM
	EDTA pH8.3	1 mM

One-Phor-All-Buffer (x10)

	Tris Acetate pH7.5	100 mM
	Magnesium Acetate	100 mM
	Potassium Acetate	500 mM

**2.1.6. Solutions for Protein work****2.1.6.1. Solutions for metal chelate chromatography**

Shown below are formulations of buffers with Tris. Where needed Tris was replaced with the appropriate buffer.

<u>Charge Buffer</u>	NiSO <sub>4</sub>	100 mM
<u>Binding Buffer</u>	Tris-HCl pH 8.0	20 mM
	NaCl	0.5 M
	Imidazole	5 mM
<u>Wash Buffer I (20 mM)</u>	Tris-HCl pH 8.0	20 mM
	NaCl	0.5 M
	Imidazole	20 mM
<u>Wash Buffer II (50 mM)</u>	Tris-HCl pH 8.0	20 mM
	NaCl	0.5 M
	Imidazole	50 mM
<u>Wash Buffer III (60 mM)</u>	Tris-HCl pH 8.0	20 mM
	NaCl	0.5 M
	Imidazole	60 mM
<u>Elute Buffer (400 mM)</u>	Tris-HCl pH 8.0	20 mM
	NaCl	0.5 M
	Imidazole	400 mM
<u>Strip Buffer</u>	EDTA	100 mM
	NaCl	0.5 M
	Tris-HCl pH 8.0	20 mM
<u>10 x His•tag cleavage buffer</u>		
	Tris-HCl pH 8.4	200 mM
	NaCl	1.5 M
	CaCl <sub>2</sub>	25 mM

Potassium Phosphate Buffer pH range: 5.8 - 8.0

Stocks of 1M K<sub>2</sub>HPO<sub>4</sub> and 1M KH<sub>2</sub>PO<sub>4</sub> were made and added together to obtain the desired pH. E.g. 9.4 ml 1M K<sub>2</sub>HPO<sub>4</sub> + 0.6 ml 1M KH<sub>2</sub>PO<sub>4</sub> made to 100 ml with dH<sub>2</sub>O gives a solution of pH 8.0. For other pH values refer to Sambrook et al.

Sodium Phosphate Buffer pH range: 5.8 – 8.0

Stocks of 1M Na<sub>2</sub>HPO<sub>4</sub> and 1M NaH<sub>2</sub>PO<sub>4</sub> were made and added together to obtain the desired pH. E.g. 9.32 ml 1M Na<sub>2</sub>HPO<sub>4</sub> + 6.8 ml 1M NaH<sub>2</sub>PO<sub>4</sub> made to 100 ml with dH<sub>2</sub>O gives a solution of pH8.0. For other pH values refer to Sambrook et al.

CHES (2-[N-Cyclohexylamino] ethane sulfonic acid) pH range: 8.6 - 10

Added to 1L of dH<sub>2</sub>O before use in protein stabilisation. Adjusted to desired pH.

MES (2-[N-Morpholino] ethane sulfonic acid pH range: 5.5 – 6.7

Added to 1L of dH<sub>2</sub>O before use in protein stabilisation. Adjusted to desired pH

HEPES pH range: 6.8 – 8.2

Added to 1L of dH<sub>2</sub>O before use in protein stabilisation

### 2.1.6.2. Solutions for SDS-PAGE and Native gels

Depending on the size of the protein and the resolution required, SDS-PAGE or Native gels were made up to the following specifications, the resolving gel being made first and then the stacking gel on top of this.

**Table 2.5: Composition of the running gels made to the desired concentration of acrylamide depending on the size of the protein**

#### Running Gels

Composites (ml)	SDS-PAGE Gels				Native Gels		
	8 %	10 %	12.5%	15 %	3 %	5 %	10 %
<b>dH<sub>2</sub>O</b>	5.7	4.7	3.4	2.2	9.7	8.7	5.2
<b>30 % Acrylamide</b>	4.0	5.0	6.3	7.5	1.52	2.55	5.0
<b>1.5 M Tris pH 8.8</b>	3.8	3.8	3.8	3.8	3.8	3.8	3.8
<b>10 % SDS</b>	1.5	1.5	1.5	1.5	-	-	-
<b>10 % APS</b>	0.15	0.15	0.15	0.15	0.15	0.15	0.15
<b>TEMED</b>	0.01	0.01	0.01	0.01	0.01	0.01	0.01

**Table 2.6: Composition of the stacking gels made to the desired concentration**

#### Stacking Gels

Composites (ml)	SDS Gel: 5 %	Native Gel: 5 %
<b>dH<sub>2</sub>O</b>	3.4	4
<b>30 % Acrylamide</b>	1.5	1.5
<b>1.5M Tris pH 6.8</b>	1.9	1.9
<b>10 % SDS</b>	0.75	-
<b>10 % APS</b>	0.075	0.075
<b>TEMED</b>	0.01	0.01

<u>10 x Running Buffer</u>	Tris-HCl	30 g/l
	Glycine	144 g/l
<u>SDS sample Buffer</u>	0.5 M Tris-HCl pH 6.8	6.00 ml
	Glycerol	4.80 ml
	10 % SDS	9.60 ml
	0.05 % bromophenol blue	1.20 ml
	dH <sub>2</sub> O	24.0 ml
	add 14 µl β-Mercaptoethanol/ml 2 x SDS-Buffer	
<u>Native sample Buffer</u>	0.5 M Tris-HCl pH 6.8	1 ml
	Glycerol	0.8 ml
	0.05 % bromophenol blue	0.2 ml
	dH <sub>2</sub> O	6 ml

Denaturing Molecular Weight Marker Dalton VII

<u>Protein</u>	<u>Molecular Weight (kDa)</u>
 Bovine serum albumin	66,000
Ovalbumin	45,000
G-3-P dehydrogenase	36,000
Carbonic anhydrase	29,000
Trypsinogen	24,000
Trypsin inhibitor	20,100
X-Lactalbumin	14,200

High molecular marker:

<u>Protein</u>	<u>Molecular Weight (kDa)</u>
myosin, rabbit muscle	205
β-galactosidase	116
phosphorylase 6	97.4
bovine albumin (BSA)	66
egg albumin	45
carbonic anhydrase	29

MWM for native electrophoresis

<u>Protein</u>	<u>Molecular Weight (Da)</u>
Thyroglobulin	669,000
Ferritin	440,000
Catalase	232,000
Lactate dehydrogenase	140,000
Bovine serum albumin	66,000

Coomassie blue stain

Trichloroacetic acid	250 ml
Coomassie blue R250	0.6 g
SDS	0.1 g
Tris-HCl	0.25 g
Glycine	0.15 g
up to 500 ml	

**2.1.6.3. Solutions for Western Blotting**

<u>10 x Running Buffer:</u>	Tris-HCl	30 g
	Glycine	144 g
	up to 1000 ml with dH <sub>2</sub> O	
<u>Transfer buffer:</u>	Methanol	200 ml
	10 x Running buffer	100 ml
	up to 1000 ml dH <sub>2</sub> O	
<u>PBS:</u>	NaCl	140 mM
	KCl	3 mM
	Na <sub>2</sub> HPO <sub>4</sub>	10 mM
	KH <sub>2</sub> PO <sub>4</sub>	2 mM
<u>Tris-NaCl buffer:</u>	Trizma base pH 7.4	50 mM
	NaCl	150 mM

Low Rainbow marker:

	<u>Protein</u>	<u>Molecular weight (kDa)</u>	<u>Colour</u>
	Ovalbumin	45	Yellow
	Carbonic anhydrase	30	Orange
	Trypsin inhibitor	20.1	Blue
	Lysozyme	14.3	Magenta
	Aprotinin	6.5	Blue black
	Insulin (b) chain	3.5	Blue
	Insulin (a) chain	2.5	Blue

High Rainbow marker:

	<u>Protein</u>	<u>Molecular weight (kDa)</u>	<u>Colour</u>
	Myosin	220	Blue
	Phosphorylase b	97	Orange
	Bovine serum albumin	66	Pink
	Ovalbumin	45	Yellow
	Carbonic anhydrase	30	Orange
	Trypsin inhibitor	20.1	Blue
	Lysozyme	14.3	Magenta

Add an appropriate volume of your gel-loading buffer (up to 10  $\mu$ l). Loading buffer should contain sufficient  $\beta$ -mercaptoethanol to give a final concentration of 5 % to ensure complete reduction of proteins. Samples may be loaded on to gel immediately, or can be stored temporarily on ice. Load 20  $\mu$ l for full-size gels (16–20 cm) and 5  $\mu$ l for mini-gels.

**2.1.6.4. Solutions for Gel Filtration Chromotography (FPLC)**Column Buffers:

HEPES Buffer pH	HEPES pH	50 mM
	NaCl	150 mM
Tris Buffer pH8	Tris HCl pH 8.0	50 mM
	NaCl	150 mM
Washing Buffer	NaCl	2 M
	Sodium Azide	0.2 %
Storage Buffer	ddH <sub>2</sub> O	
	Sodium Azide	0.2 %

## 2.2. Molecular biology methods

Protocols are based on the methods described by Sambrook et al (2002).

### 2.2.1. Isolation of genomic DNA from PAO

*Ps. aeruginosa* genomic DNA was kindly received from Dr Evelyne Raux-Deery (University of Kent), at a concentration of 0.02 mg/ml. *Paracoccus denitrificans* genomic DNA was kindly donated by Richard Zajicek (University of Oxford).

### 2.2.2. Isolation of plasmid DNA

Plasmid DNA was isolated from bacterial strains such as JM109 or SCS110, using the Qiagen QIAprep<sup>®</sup> Miniprep kit and the protocol for micro-centrifuges. Cultures were grown overnight in 5 ml LB and antibiotics where appropriate and then pelleted by centrifugation before isolation of the DNA. The DNA was eluted into a sterile tube and stored at  $-20^{\circ}\text{C}$  until further use.

**Table 2.7: Amino acids listed with their properties**

**Stop codons include: UAA UAG UGA**

Amino Acid	Abbreviations		Codons	Group
Alanine	Ala	A	GCA GCC GCG GCU	Small
Cysteine	Cys	C	UGC UGU	Nucleophilic
Aspartic Acid	Asp	D	GAC GAU	Acidic
Glutamic Acid	Glu	E	GAA GAG	Acidic
Phenylalanine	Phe	F	UUC UUU	Aromatic
Glycine	Gly	G	GGA GGC GGG GGU	Small
Histidine	His	H	CAC CAU	Basic
Isoleucine	Ile	I	AUA AUC AUU	Hydrophobic
Lysine	Lys	K	AAA AAG	Basic
Leucine	Leu	L	UUA UUG CUA CUC CUG CUU	Hydrophobic
Methionine	Met	M	AUG	Hydrophobic
Asparagine	Asn	N	AAC AAU	Amide
Proline	Pro	P	CCA CCC CCG CCU	Hydrophobic
Glutamine	Gln	Q	CAA CAG	Amide
Arginine	Arg	R	AGA AGG CGA CGC CGG CGU	Basic
Serine	Ser	S	AGC AGU UCA UCC UCG UCU	Nucleophilic
Threonine	Thr	T	ACA ACC ACG ACU	Nucleophilic
Valine	Val	V	GUA GUC GUG GUU	Hydrophobic
Tryptophan	Trp	W	UGG	Aromatic
Tyrosine	Tyr	Y	UAC UAU	Aromatic

### 2.2.3. Restriction of DNA

The amplified DNA was incubated for a minimum of 2 hours with the restriction enzymes in the corresponding buffers, at the temperature required by the restriction enzyme. The fragments were then separated on an agarose gel (Section 2.2.4.1) and if desired ligated as needed (Section 2.2.6). Please refer to the Promega restriction enzyme (RE) table for double digests. For the majority of double digests, OPAB buffer was sufficient.

Typical restriction digests:

	Vector	5' RE	3'RE	ddH <sub>2</sub> O	Total Vol:
pGEMT-E	5 µl	1 µl	1 µl	3 µl	10 µl
pET14b	15 µl	1 µl	1 µl	3 µl	20 µl
pUCP19	4 / 5 µl	1 µl	1 µl	3 / 4 µl	10 µl

### 2.2.4. Electrophoresis of DNA

DNA fragments were separated, identified and purified by agarose gel electrophoresis (1 % w/v agarose in TAE buffer). Linear fragments of DNA migrate through the gel according to their molecular weight and charge.

#### 2.2.4.1. Agarose gel

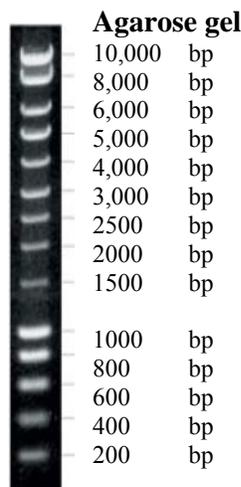
In this project 1 % agarose gels were sufficient to separate the DNA fragments according to the size of the fragments. Agarose (1 g) was dissolved in 1 x TAE Buffer (100 ml) with the addition of ethidium bromide. Ethidium bromide is a fluorescent dye that intercalates between the base pairs, allowing visualization of the DNA under UV light. DNA samples in 20 % DNA loading buffer were loaded into the molded gel wells and a voltage of 70 mV (at most 90 mV) was applied to separate the DNA. Hyperladder I (Section 2.2.4.3) was also applied to determine the size of the separated fragments.

#### 2.2.4.2. Visualisation and UV

Ethidium bromide forms a complex with DNA and absorbs UV radiation (312 nm), and re-emits it at 590 nm in the red-orange wavelength range. This allows the DNA fragments to be visualized by placing the gel onto an UV transilluminator. A camera then photographed the gel through a red filter.

### 2.2.4.3. Hyperladder (Bioline)

Molecular weight marker Hyperladder I (Bioline)



### 2.2.5. Isolation and purification of a DNA fragment from an Agarose gel

DNA bands of the correct size were carefully excised (separately) from an agarose gel using separate sterile scalpel blades. Each slice was then purified using the Qiagen QIAquick<sup>®</sup> gel extraction kit and QIAquick<sup>®</sup> gel extraction kit protocol for micro-centrifuges in QIAquick<sup>®</sup> spin Handbook. The DNA was eluted into a sterile tube and stored at  $-20\text{ }^{\circ}\text{C}$  until further use.

### 2.2.6. Ligation of DNA fragments

After restriction and purification of inserts or plasmid / vectors, DNA fragments were ligated together using T4 DNA Ligase in 1 x T4 DNA Ligase Buffer (Promega). The reaction was allowed to proceed usually at  $4\text{ }^{\circ}\text{C}$  overnight or at  $16\text{ }^{\circ}\text{C}$  for 2 hrs. For optimum results a ratio of 1:1.5 of insert to vector was used.

## **2.2.7. Competent cells and Transformation**

### **2.2.7.1. Preparation of *E. coli* Competent Cells**

An individual colony from an overnight plated culture of the desired bacterium was taken and used to inoculate 50 ml of LB and then incubated at 37 °C with aeration. The culture was grown to an OD<sub>600</sub> of approximately 0.3. After standing the culture on ice for 10 min to cool the cells, the cells were harvested by centrifugation at 3000 rpm for 10 min at 4 °C. The cell pellet was washed with 50 ml ice cold 0.1 M MgCl<sub>2</sub>, centrifuged again and resuspended in 25 ml ice cold 0.1 M CaCl<sub>2</sub> and left on ice for 20 min to 1 hr. The cells were centrifuged again and the pellet was resuspended in 250 µl ice-cold 0.1 M CaCl<sub>2</sub> / 15 % glycerol. The final bacterial suspension was dispensed into 1.5 ml eppendorfs in 20 µl aliquots and immediately frozen at -80 °C.

### **2.2.7.2. Transformation of *E. coli* Competent Cells**

The method is based on the heat shock transformation protocol. Competent cells were defrosted on ice for 15 min before adding of 0.5 µl to 1.5 µl of miniprep or ligation mixture. After a further incubation on ice for 10 min, the cell-DNA mixture was heat shocked for 50 sec at 42 °C and cooled on ice for 2 min before adding of 200 µl LB medium. The cell-DNA mixture was then incubated at 37 °C for 20 min 1 hr to allow plasmid replication and for cells to confer antibiotic resistance that allows selection of the cells containing the plasmid when plated. The cells were then streaked out onto LB plates with appropriate antibiotics and incubated overnight at 37 °C.

### **2.2.7.3. Preparation of *Ps. aeruginosa* Competent cells**

The protocol described here is adapted from the method described by Chuanchuen, Narasaki et al. 2002. Wild type *Ps. aeruginosa* cells were grown overnight at 37 °C from a single colony in 10 ml of LB. For harvesting, 1.5 ml aliquots of cells were transferred to pre-chilled microfuge tubes and the cells pelleted by centrifugation at 13000-x g for 30 secs. The cell pellets were resuspended in 1 ml cold (4 °C) 0.1 M MgCl<sub>2</sub> using a micropipette, centrifuged again at 13000 x g for 30 secs and the supernatant discarded. The pellets were then resuspended in 1 ml transformation salts with glycerol (TG Salts, see 2.1.4) and incubated on ice for 10 min. The cell pellets were centrifuged at 13000 x g for 30 sec and the supernatant decanted. The

resultant pellets were resuspended in 200 µl cold TG salt solution, divided into 100 µl aliquots into prechilled eppendorfs and frozen at -80 °C for later use

#### **2.2.7.4. Transformation of *Ps. aeruginosa* Competent cells**

This method has also been adapted from the method described by (Chuanchuen, Narasaki et al. 2002). Competent cells were defrosted ice for 15 min before the addition of 1.5 µl of miniprep. After a further incubation on ice for 15 min, the cell-DNA mixture was heat shocked for either 50 sec at 42 °C or 2 min at 37 °C and then cooled on ice for 2 min before the addition of 500 µl LB medium. The cell-DNA mixture was then incubated at 37 °C for no less than 1 hr, with or without shaking, to allow plasmid replication and for the cells to confer antibiotic resistance that allows selection of the cells containing the plasmid when plated. The cells were then streaked out onto LB plates containing the appropriate antibiotics and incubated at 37 °C for 36 - 48 hrs.

#### **2.2.7.5. Preparation of *S. enterica* AR3612 Competent cells**

*S. enterica cysG metE* (AR3612) cells were kindly provided by Dr Evelyne Raux-Deery. These cells are deficient in CysG<sup>A</sup> activity but retain their CysG<sup>B</sup> activity as the host contains a gene having the same phenotype as *cysG<sup>B</sup>* (*cbiK*) (Raux, Lanois et al. 1996). Cells were streaked over night on LB media with chloramphenicol, and an individual colony was taken to inoculate 50 ml of LB with chloramphenicol 34 mg/ml and incubated at 37 °C with aeration. The culture was grown to an OD<sub>600</sub> of approximately 0.3. After standing the culture on ice for 10 min to cool the cells, the cells were harvested by centrifugation at 3000 rpm for 10 min at 4 °C. The cell pellet was washed with 50 ml ice cold 0.1 M MgCl<sub>2</sub>, centrifuged again and resuspended in 25 ml ice cold 0.1 M CaCl<sub>2</sub> and left on ice for 20 min to 1 hr. The cells were centrifuged again and the pellet was resuspended in 250 µl ice-cold 0.1 M CaCl<sub>2</sub> / 15 % glycerol. The final bacterial suspension was dispensed into 1.5 ml eppendorfs in 50 µl aliquots and immediately frozen at -80 °C.

### 2.2.7.6. Transformation of *S. enterica* AR3612 Competent cells

Competent cells were defrosted on ice for 15 min before the addition of 5  $\mu$ l of miniprep as desired. After a further incubation on ice for 10 min, the cell-DNA mixture was heat shocked for 50 sec at 42 °C and then cooled on ice for 2 min before the addition of 200  $\mu$ l LB medium. The cell-DNA mixture was incubated at 37 °C for no less than 30 min to 1 hr allowing plasmid replication and for the cells to confer antibiotic resistance allowing selection of the cells when plated. The cells were then streaked out onto LB plates containing the appropriate antibiotics before incubating at 37 °C for 24 – 48 hrs until colonies are formed.

### 2.2.7.7. *S. enterica* *cysG* complementation

A *S. enterica* *cysG metE* (AR3612) strain is deficient in its ability to transform UIII into PC-2. Complementation studies were conducted by transforming 5  $\mu$ l of the vector miniprep containing a plasmid of the gene or combination of genes into *S. enterica* *cysG metE* (AR3612) strain competent cells. The resultant colonies were then restreaked on to minimal media plates made with 50  $\mu$ g/ml of both methionine and cysteine; with cysteine and without methionine; without cysteine and with methionine; without cysteine and methionine; plus ampicillin 100 mg/l and Chloramphenicol 34 mg/ml as appropriate. The plates were then incubated at 37 °C for 24 – 48 hrs, until bacterial colonies were evident in the plates.

### 2.2.7.8. Glycerol stocks of strains

Overnight bacterial cultures of all the various strains (Table 2.2), were grown in 5ml LB plus or minus the appropriate antibiotic, and prepared for storage by the addition of 15 % glycerol (volume/volume) to create a library of strains. The strains were stored at –80 °C.

A typical stock would contain:

Overnight culture	850 $\mu$ l
100 % Glycerol	<u>150 <math>\mu</math>l</u>
	1000 $\mu$ l

### 2.2.8. Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) is the process of amplification of the DNA of the gene of interest. PCR amplification involves the introduction of restrictable enzymes, which in turn allows cloning into pre-restricted plasmids or vectors. All PCR reactions were prepared on ice as below and performed in a Techne thermal cycler with cycles as indicated in Section 2.2.8.1. Following PCR, reaction products are separated by an agarose gel (Section 2.2.4.) and the band of correct size and interest is cut from the gel. The DNA of the gene of interest is then purified from the gel by gel extraction (Section 2.2.5).

<u>PCR Reaction mixture:</u>	+ Q Buffer	- Q Buffer
10 x PCR buffer (Promega)	5 $\mu$ l	5 $\mu$ l
5 x Q Buffer (Promega)	10 $\mu$ l	10 $\mu$ l
2mM dNTPs (Bioline)	5 $\mu$ l	5 $\mu$ l
Genomic DNA (0.02 mg/ml)	1 $\mu$ l	1 $\mu$ l
10 $\mu$ M 5' Primer	2 $\mu$ l	2 $\mu$ l
10 $\mu$ M 3' Primer	2 $\mu$ l	2 $\mu$ l
Sterile water	24.5 $\mu$ l	34.5 $\mu$ l
Hot Star Taq (Promega)	<u>0.5 <math>\mu</math>l</u>	<u>0.5 <math>\mu</math>l</u>
	50 $\mu$ l	50 $\mu$ l

#### 2.2.8.1. PCR of *nir* genes

All the *nir* genes were amplified by PCR using the respective designed primers (Table 2.8).

##### 2.2.8.1.1. *Nde*I and *Bam*HI Primers

Stated are the reaction conditions for the genes cloned with the *Nde*I and *Bam*HI primers, for single gene plasmids.

<i>nirE</i> :	1 cycle:	95 °C for 15 minutes
	35 cycles:	{ 95 °C for 20 seconds 60 °C for 20 seconds 72 °C for 51 seconds
	1 cycle:	72 °C for 10 minutes

***nirD, nirL, nirG, nirH:***

1 cycle: 95 °C for 15 minutes

35 cycles: { 95 °C for 20 seconds  
60 °C for 20 seconds  
72 °C for 35 seconds

1 cycle: 72 °C for 10 minutes

***nirF***

1 cycle: 95 °C for 3 minutes

35 cycles: { 95 °C for 20 seconds  
55 °C for 20 seconds  
72 °C for 2 minutes 20 seconds

1 cycle: 72° C for 10 minutes

**2.2.8.1.2. *NdeI* or *NheI*, RBS and *XbaI* and *BamHI* Primers**

Below are the reaction conditions for the genes cloned with the ribosome-binding site (RBS) and *NheI* or *NdeI* primers at the 5' end and *XbaI* and *BamHI* primers at the 3' end, for multiple gene plasmids.

1 cycle: 95 °C for 3 minutes

35 cycles: { 95 °C for 20 seconds  
55 °C for 20 seconds  
72 °C for 2 minutes 20 seconds

1 cycle: 72°C for 10 minutes

### 2.2.8.2. Primers for PCR

**Table 2.8: Tabulation of all primers used in this study to clone the genes individually and for the Link and Lock method.**

PCR primers were designed for each of the genes from a database nucleotide sequence, Pubmed Acc # D84475, for genes from *Ps.aeruginosa*, unless otherwise indicated.

**NdeI = CATATG**      **BamHI = GGATCC**      **Stop = TAG/TAA/TGA**

**XbaI = TCTAGA**      **SpeI = ACTAGT**      **NheI = GCTAGC**

**RBS (ribosome binding site) in pET14b = AGGAG (A)**

Gene of Interest	NdeI and BamHI Primers
<i>nirF</i> (PAO) 5'- 3' NdeI	TGC ATA TGA ACC TGC GTC C
<i>nirF</i> (PAO) 3'- 5' BamHI	TCG GAT CCT CCT AGA GTC CGA TGT GC
<i>nirD</i> (PAO) 5'- 3' NdeI	TTC ATA TGG ACG ACC TCT CC
<i>nirD</i> (PAO) 3'- 5' BamHI	GTG GAT CCT CAT CGG CTG GCC TCC AG
<i>nirL</i> (PAO) 5'- 3' NdeI	TGC ATA TGA ACC CCG TCG AAC C
<i>nirL</i> (PAO) 3'- 5' BamHI	TTG GAT CCT CAT CCA TGG CGG CGC TCC
<i>nirG</i> (PAO) 5'- 3' NdeI	TGC ATA TGG ATG AGT TCG ACC
<i>nirG</i> (PAO) 3'- 5' BamHI	TGG GAT CCG TCA GAG GGG GAA ATG C
<i>nirH</i> (PAO) 5'- 3' NdeI	TCC ATA TGT CGG CCT GCA TTT CC
<i>nirH</i> (PAO) 3'- 5' BamHI	GTG GAT CCT CAG CAT GGA CGT TCCTCC
<i>nirE</i> (PAO) 5'- 3' NdeI	TGC ATA TGA ACA CTA CCG TGA TTC CG
<i>nirE</i> (PAO) 3'- 5' BamHI	TCG GAT CCG ATC AGG CGC ATG CGA C
<b>Link and Lock Primers</b>	
<i>nirF</i> (PAO) 5'- 3' NheI-RBS	TCA GCT AGC GAG GAG TTA TAT ATC ATG AAC CTG CGT CCC CTC
<i>nirF</i> (PAO) 3'- 5' XbaI-BamHI	TAT GGA TCC AAA TCT AGA TCC TAG AGT CCG ATG TGC
<i>nirD</i> (PAO) 5'- 3' NheI-RBS	ATC GCT AGC AGG AGA TCA CAG CTC ATG GAC GAC CTC TCC CGC
<i>nirD</i> (PAO) 3'- 5' XbaI-BamHI	TAT GGA TCC AAA TCT AGA TCA TCG GCT GGC CTC CAG
<i>nirL</i> (PAO) 5'- 3' NheI-RBS	GAT GCT AGC AGG AGA TCA CAC CAG ATG AAC CCC GTC GAA CC
<i>nirL</i> (PAO) 3'- 5' XbaI-BamHI	GTA GGA TCC AAA TCT AGA TCA TCC ATG GCG GCG CTC C
<i>nirG</i> (PAO) 5'- 3' NheI-RBS	TAT GCT AGC AGG AGA TCA CAG ATC ATG GAT GAG TTC GAC CGG

<i>nirG</i> (PAO) 3'- 5' <i>XbaI</i> - <i>Bam</i> HI	GAA <b>GGA TCC</b> AGA <b>TCT AGA</b> GTC AGA GGG GGA AAT GCA G
<i>nirH</i> (PAO) 5'- 3' <i>NheI</i> -RBS	CTA <b>GCT AGC AGG</b> AGA TCT CAG CTC ATG TCG GCC TGC ATT TCC
<i>nirH</i> (PAO) 3'- 5' <i>XbaI</i> - <i>Bam</i> HI	TCA <b>GGA TCC</b> TAA <b>TCT AGA</b> TCA GCA TGG ACG TTC CTC C
<i>nirJ</i> (PAO) 5'- 3' <i>NheI</i> -RBS	TAT <b>GCT AGC AGG</b> AGA AGA TTG TCC ATG CTG AGG ATC AG
<i>nirJ</i> (PAO) 3'- 5' <i>XbaI</i> - <i>Bam</i> HI	TTA <b>GGA TCC</b> AAA <b>TCT AGA</b> CCT TAG ACG GCG TGC AGC
<i>nirE</i> (PAO) 5'- 3' <i>NheI</i> -RBS	TCA <b>GCT AGC GAG GAG</b> TTA TAT ATC ATG AAC ACT ACC GTG ATT CCG
<i>nirE</i> (PAO) 3'- 5' <i>XbaI</i> - <i>Bam</i> HI	TAT <b>GGA TCC</b> AAA <b>TCT AGA</b> GAT CAG GCG CAT GCG AC
<i>nirJ</i> ( <i>Pa. Pantotrophus</i> ) 5'- 3' <i>NdeI</i>	TAT <b>CAT ATG</b> TTC CGC CTG ACC C
<i>nirJ</i> ( <i>Pa. Pantotrophus</i> ) 5'- 3' <i>NheI</i> -RBS	ATC <b>GCT AGC AGG</b> AGA TCA CAG CTC ATG TTC CGC CTG ACC C
<i>nirJ</i> ( <i>Pa. Pantotrophus</i> ) 3'- 5' <i>XbaI</i> - <i>SacI</i>	ATA GAG CTC AGA <b>TCT AGA</b> CTA AAG AAA GCG GTG GGC

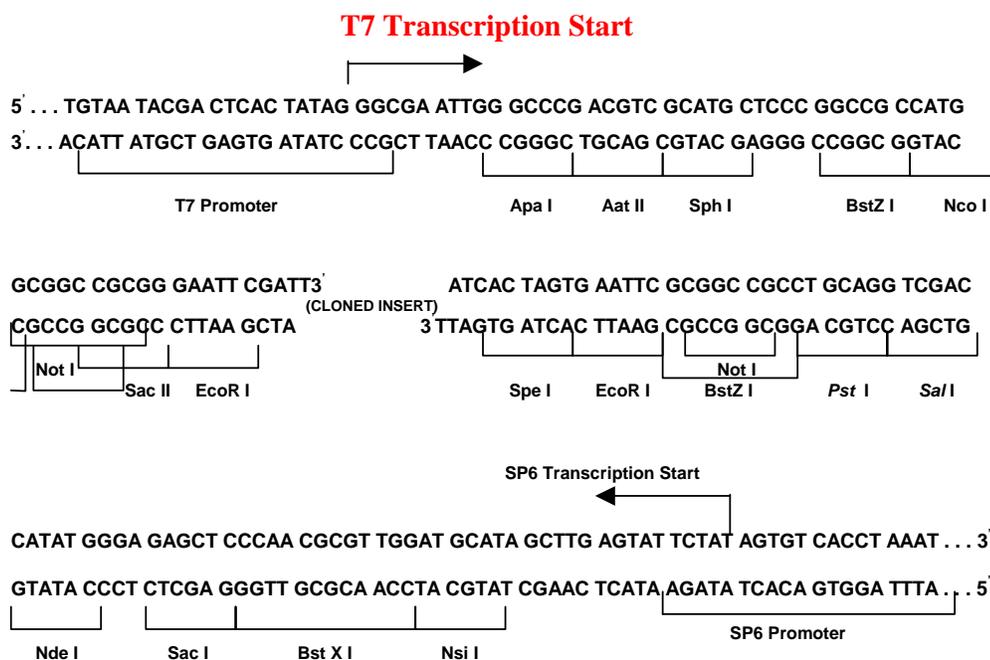
**NdeI = CATATG****BamHI = GGATCC****Stop = TAG/TAA/TGA****XbaI = TCTAGA****SpeI = ACTAGT****NheI = GCTAGC****RBS in pET14b = AGGAG (A)**

### 2.2.9. T-Vector cloning of PCR products

All PCR products of predicted size, having been isolated and purified from an agarose gel (Section 2.2.5) were cloned using the pGEM<sup>®</sup>-T Easy vector system from Promega (refer to the manual more in depth details or plasmid map in the appendix). Briefly, an 'A'- overhang generated on the DNA fragment during PCR can be ligated to the cohesive 'T'- overhang of pGEM-T Easy without the need for restriction prior to ligation Figure 2.1. The ligation mixture is then transformed in JM109 cells (Section 2.2.7.2) incubated overnight on IPTG / X-gal / Amp plates for blue/white selection. Colonies that were white were selected as they contained the plasmid with the gene of interest, due to the disruption of the  $\beta$ -galactosidase gene. Those that were blue were not selected, as they were re-ligated plasmid

**Figure 2.1: The promoter and multiple cloning sequences of pGEMT and pGEMT-Easy Vectors.**

The top strand of the sequence shown corresponds to the RNA synthesised by T7 RNA polymerase. The bottom strand corresponds to the RNA synthesised by SP6 RNA polymerase.



**2.2.10. Subcloning into pET14b:**

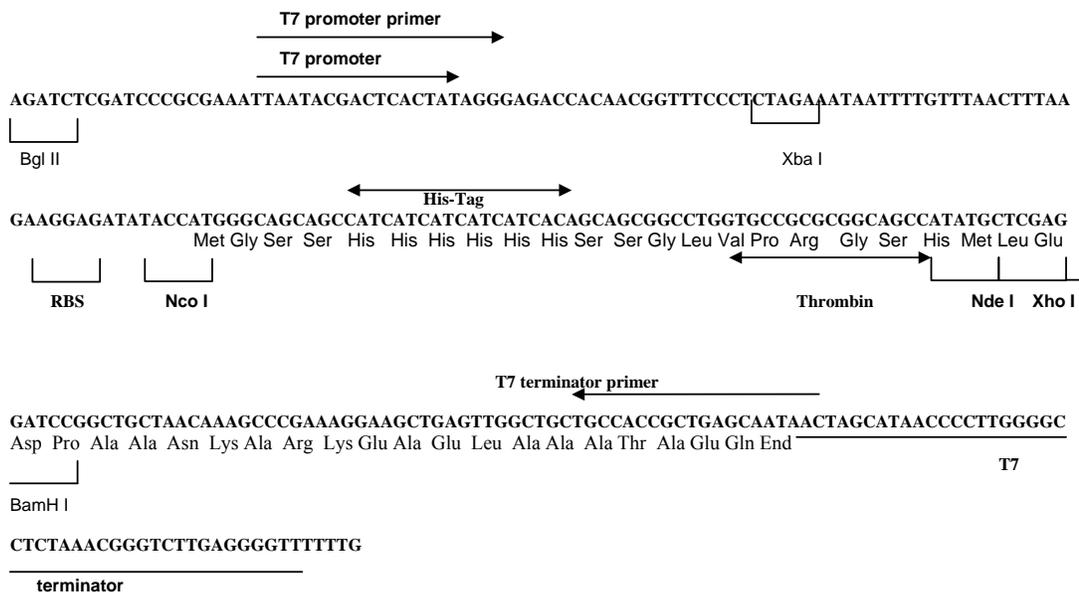
**NdeI – BamHI for Protein Expression in pET14b**

The genes with primers designed to introduce *NdeI* (CATATG) and *BamHI* (GGATCC) at the 5' and 3' sites respectively, were all subcloned into pET14b, via pGEM-T<sup>®</sup> E vector containing the PCR product. The introduction of these sites allowed easy in frame cloning into pET14b, which contains a N-terminal histag sequence (CAT x 6) allowing the encoded protein to be produced as a N-terminal poly-histidine tagged fusion protein. This conveniently allows protein overproduction and recombinant protein affinity purification by metal chelate chromatography. For the full method refer to the Novagen pET system instruction manual (Novagen). To confirm the sequence of the gene was correct, all vectors were sent off for sequencing.

The gene and plasmid were restricted to have complementary sites (Section 2.2.3) and ligated accordingly (Section 2.2.6) after isolating the fragment by agarose gel (Section 2.2.4.) and gel extraction (Section 2.2.5).

**Figure 2.2: The promoter and multiple cloning sites of pET14b**

The pET-14b vector carries an N-terminal HisTag sequence followed by a thrombin site and three cloning sites. The cloning / expression region of the coding strand transcribed by T7 RNA polymerase is shown below.

**2.2.11. Subcloning into pUCP Nde/Nco for Expression in *Ps. aeruginosa***

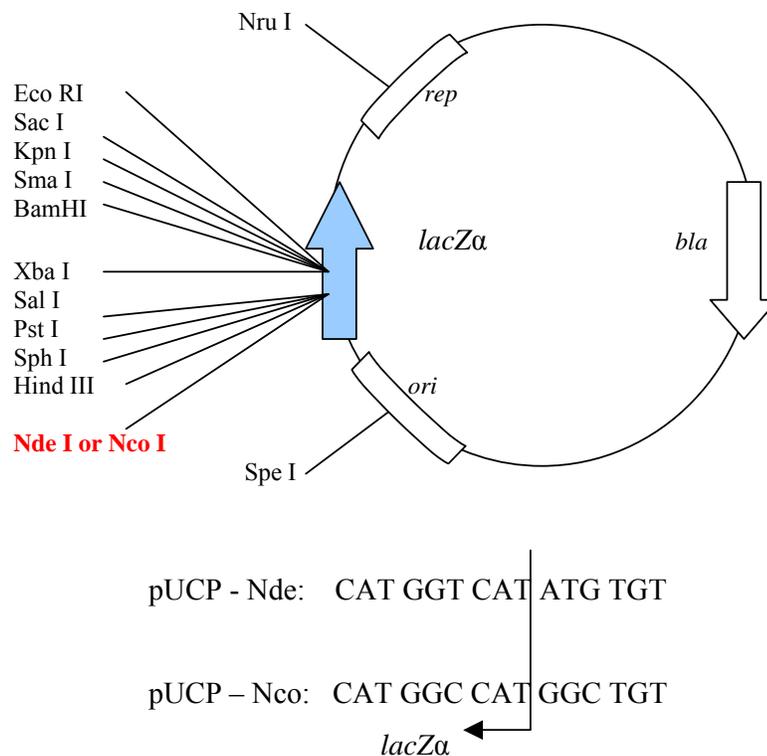
Since *Ps. aeruginosa* does not recognise the T7 promoter, a new system was formulated to facilitate recombinant protein overproduction in *Ps. aeruginosa*. Various plasmids were screened for transformation efficiency in *Ps. aeruginosa* and suitable multiple cloning sites. The plasmid most compatible in both cases was pUCP19-Nde / Nco an *Escherichia-Pseudomonas* shuttle vector for recombinant protein over expression in *Pseudomonas* (Cronin and McIntire 1999). This vector, being under the control of a *lac* promoter, allows the genes to be over expressed and induced to over produce the proteins by the addition of IPTG. A simplified map of the vector is shown in Figure 2.3, please refer to the paper for more detail (Cronin and McIntire 1999).

The *nir* genes were individually subcloned from pET14b into pUCP19-NcoI ( $\text{amp}^R$  carb<sup>R</sup>), using the pET14b restriction sites *NcoI* (CCATGG) *BamHI* (GGATCC). This allowed both the gene and the histag to be subcloned into pUCP19-NcoI and allowed recombinant protein overproduction in *Ps. aeruginosa*. The resultant

homologously expressed recombinant protein was then purified by metal chelate chromatography.

Similarly, the genes were all subcloned using the *Nde*I (CATATG) and BamHI (GGATCC) restriction sites, into pUCP19-*Nde* ( $\text{amp}^R$   $\text{carb}^R$ ). This allowed expression of the various genes without encoding for a histag, subsequently producing a non-histagged recombinant protein in *Ps. aeruginosa*.

**Figure 2.3: The promoter and multiple cloning sites of pUCP-Nde / pUCP-Nco**  
Protein expression vectors pUCP-Nco and pUCP-Nde. The relative lengths and coding directions of the *lacZ $\alpha$* ,  $\beta$ -lactamase (*bla*), and pRO1614-derived replication (*rep*) genes within the pUCP-Nde vector are indicated, together with the ColE1 replication origin (*ori*) and unique restriction sites. The differences between the pUCP-Nde and pUCP-Nco vectors are four base pairs surrounding the *lacZ $\alpha$*  translational start codon that places it within the context of, respectively, a *Nde*I or *Nco*I restriction enzyme cloning site (Cronin and McIntire 1999).



### 2.2.12. Link and Lock construction of haem $d_1$ operon

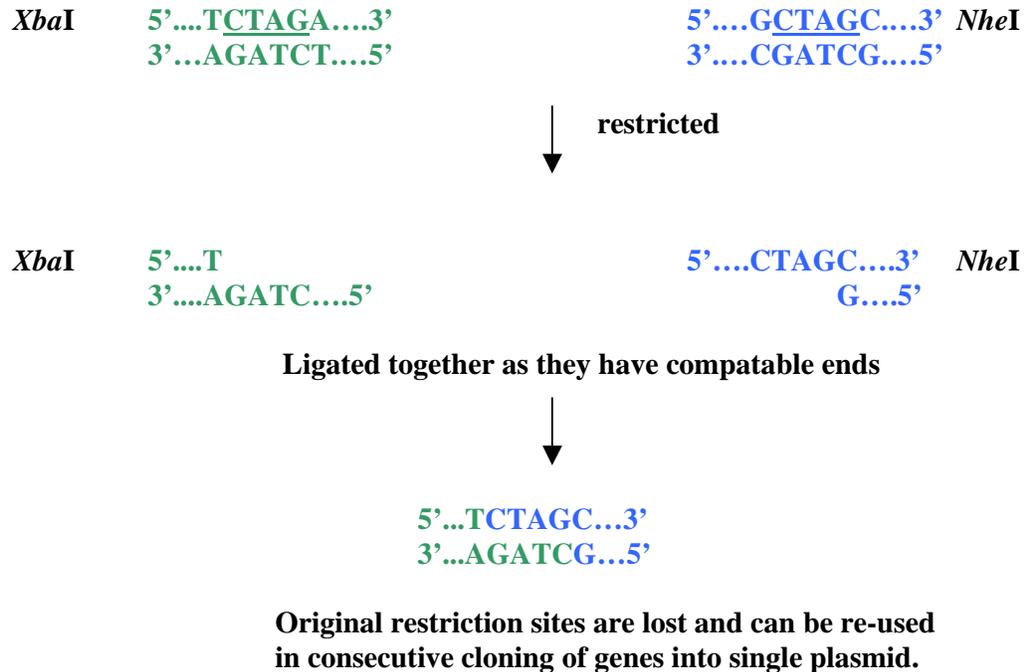
The Link and Lock Cloning strategy, as devised by Dr Evelyne Deery and Dr Helen McGoldrick, allows the consecutive cloning of genes into a plasmid in a specified order. It utilises the ability of restriction enzymes to restrict at different recognition sequences but generating the same sticky end overhangs, thus forming compatible cohesive ends. These when ligated, do not generate the same restriction sites and therefore can be re-used in introducing other genes, under the proviso that they do not occur naturally within the genes or plasmid. In this way we utilise the ability of neoisoschizomer restriction enzymes to clone our operon.

In this study we used *NheI* (GCTAGC) and *XbaI* (TCTAGA) both forming compatible cohesive ends, which when ligated are no longer restrictable, see Figure 2.4. These sites are not found within the *nir* genes or the plasmid, pUCPNde/Nco. Primers were designed for each of the genes *nirE*, *nirJ*, *nirG*, *nirH*, *nirL*, *nirD* and *nirF*, to introduce, *NdeI* (CATATG) and a ribosome binding site (AGGAG (A)) for the first gene (*nirE* or *nirF*) and then *NheI* (GCTAGC) sites and a ribosome binding sites (AGGAG (A)) at the 5' end for all other genes. *XbaI* (TCTAGA) and *BamHI* (GGATCC) restriction sites were introduced at the 3' termini of all the *nir* genes. Refer to Table 2.8 for a comprehensive list of primers. The introduction of these sites allowed link and lock cloning of the operon into the pUCPNde expression vector allowing the encoded proteins to be expressed both within *E. coli* (JM109) and wild type *Ps. aeruginosa* (PAO1).

The first gene in the sequence, for example *nirE*, was cloned into the plasmid using *NdeI* at the 5' end and *XbaI* – *BamHI* at the 3' end. The next gene and all subsequent genes are then cloned with *NheI* at the 5' end and *XbaI* – *BamHI* at the 3' end. The plasmid is then restricted with *XbaI* and *BamHI*, and the gene to be consecutively cloned is restricted with *NheI* and *BamHI*. These are then ligated utilising the compatibility of *XbaI* and *NheI*, and *BamHI* to *BamHI*, to form a plasmid with two genes. Since ligated *XbaI* and *NheI* are no longer restrictable, they can be reused to consecutively clone genes into the plasmid in subsequent steps. This has been shown diagrammatically in Figure 2.5.

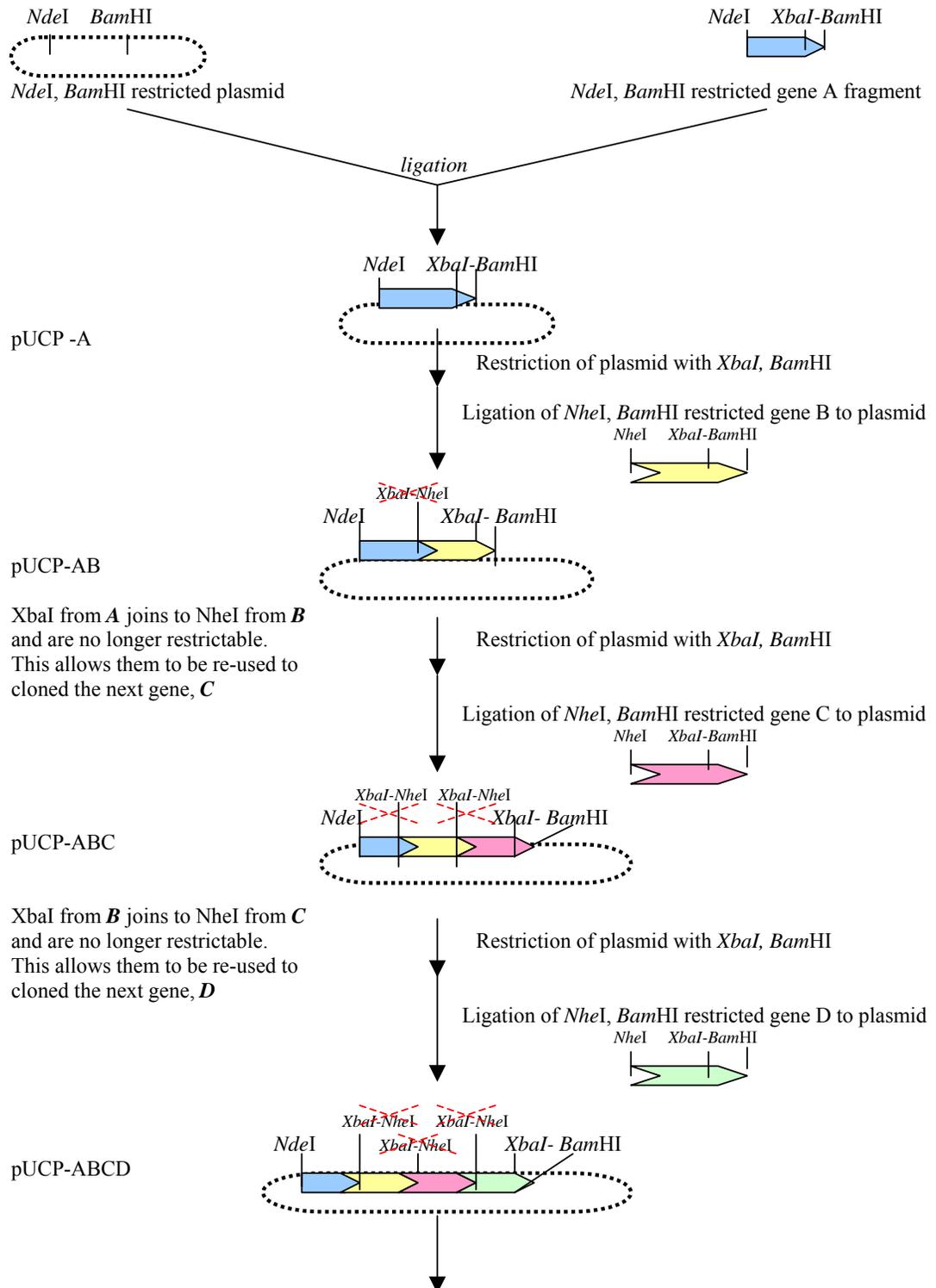
**Figure 2.4: *Xba*I and *Nhe*I; the formation of compatible cohesive ends**

*Xba*I and *Nhe*I are examples of neoschizomers with a clear advantage. Both recognize the same sequence but *Xba*I cleaves after the first T, while *Nhe*I cleaves after the first G generating the same overhang and compatible cohesive ends. These when ligated are no longer restrictable by *Xba*I and *Nhe*I, allowing their subsequent reuse. In this case only *Bfa*I can recleave the same site.



**Figure 2.5: Diagrammatic representation of link and lock method of cloning**

*XbaI* and *NheI* have been used to illustrate this but other combinations are possible.



**In this manner we may continue to clone genes consecutively into the plasmid, to obtain operons in a single plasmid**

### 2.2.13. Restriction analysis of *nir* genes in haem *d*<sub>1</sub> operon

To aid in the detection of each *nir* gene as it was consecutively cloned into the plasmid, restriction enzymes unique to each gene were identified for ease of restriction analysis. This has been summarised in Table 2.9.

**Table 2.9: Restriction enzymes unique to each *nir* gene and plasmid**

Each *nir* gene is shown relative to its position in the original operon, thus each restriction enzyme site is shown relative to its position in the gene, and not in the plasmid. The positions of restriction enzymes cutting once or twice in the plasmid are shown also. \* = in MCS

	<b>PUCPNde/ Nco</b>	<b>nirF</b> 391 – 1569	<b>nirD</b> 1578 – 2030	<b>nirL</b> 2027 – 2551	<b>nirG</b> 2544 – 2987	<b>nirH</b> 2962 - 3477	<b>nirJ</b> 3471 - 4634	<b>nirE</b> 4645 - 5484
<b>AatII</b>	<b>424 / 3831</b>	<b>1256</b>	X	<b>2268</b>	X	<b>3206</b>	X	X
<b>AvaI</b>	<b>1626</b>	<b>842 / 868 / 1521</b>	<b>1794</b>	X	X	X	<b>4313 / 4605</b>	X
<b>AscI</b>	X	X	X	X	X	X	<b>4605</b>	X
<b>BamHI</b>	<b>1631</b>	X	X	X	X	X	X	X
<b>EcoRI</b>	<b>1610</b>	<b>1411</b>	X	X	X	X	X	X
<b>HincII</b>	<b>844 / 1643*</b>	<b>861 / 1311</b>	<b>1847</b>	<b>2284</b>	X	X	X	<b>5016</b>
<b>NarI</b>	<b>196</b>	X	<b>2139</b>	<b>2561</b>	X	<b>3276</b>	X	<b>4715 / 4979 / 5318 / 5441</b>
<b>NcoI</b>	<b>1681 if pUCPNco</b>	X	X	<b>2542</b>	X	<b>3317</b>	X	X
<b>NdeI</b>	<b>1681 if pUCPNde</b>	X	X	X	X	X	X	X
<b>NheI</b>	X	X	X	X	X	X	X	X
<b>NruI</b>	<b>1353</b>	<b>626</b>	X	X	<b>2833</b>	X	X	X
<b>PstI</b>	X	X	X	X	X	X	<b>4491</b>	<b>5132</b>
<b>SalI</b>	X	<b>859</b>	X	<b>2282</b>	X	X	X	<b>5014</b>
<b>SpeI</b>	<b>2020</b>	X	X	X	X	X	X	X
<b>XbaI</b>	<b>1637*</b>	X	X	X	X	X	X	X
<b>XhoI</b>	X	X	X	X	X	X	<b>4605</b>	X

### **2.3. Biochemical Methods**

#### **2.3.1 Protein assay (Bradford Assay)**

The Bradford assay allows the estimation of the quantity of protein in a solution. Briefly, a 5 µl of the sample aliquot of eluted protein was added to 5 µl of coomassie blue G250 dye (BioRad) and a 15 µl of dH<sub>2</sub>O. The dye has absorbance maxima between 470 to 650 nm in the red cationic state, but in the presence of protein this shifts to the blue anionic form, having absorbance maxima of 590 to 620 nm. Thus, allowing the exact quantity of protein to be determined by observation at 595 nm, referring to a standard curve generated using bovine serum albumin at 5, 10, 15, 20 and 25 µg.

#### **2.3.2 A<sub>280</sub> protein concentration estimation**

Recording the optical density of the protein solution at 280 nm and the calculated extinction co-efficient for the protein, and then using the following equation, determined the concentration of protein in a solution (Layne 1957):

$$A_{280} = \text{concentration} \times \text{extinction co-efficient}$$

Further, to obtain mg/ml:

$$\text{mg / ml protein concentration} = \text{molecular mass} \times \text{concentration}$$

The extinction co-efficients and molecular weights have been recorded in the appendix for ease of reference.

### 2.3.3 Standard cloning, growth and purification of histagged proteins

Detailed information on growth parameters for each of the proteins overproduced can be found in the respective chapters.

#### 2.3.3.1 Protein overproduction in *E. coli*

To facilitate recombinant protein overproduction, the pET14b-*nir* gene plasmid constructs were individually transformed into *E. coli* BL21 (DE3) pLysS strains in which the T7 polymerase is induced by the addition of IPTG (Novagen).

An overnight starter culture of 1 colony in 10 ml LB (plus ampicillin 100 mg/l and Chloramphenicol 34 mg/l) grown aerobically at 37 °C, was used to inoculate 1L LB medium and grown at 37 °C plus ampicillin 100 mg/l and Chloramphenicol 34 mg/l at the 37 °C, until the OD<sub>600</sub> reached approximately 0.6. IPTG was then added to a final concentration of 0.4 mM. Cultures were further incubated as appropriate at 37 °C or 16 °C, before harvesting by centrifugation (10 000 x g, 4 °C). The bacterial pellets were resuspended in the appropriate binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM buffer). Proteins were purified as described in the pET system manual (Novagen) using metal chelate chromatography.

#### 2.3.3.2 Protein overproduction in *Ps. aeruginosa*

The appropriate pUCP – *nir* histidine tagged plasmid constructs were individually transformed into *Ps. aeruginosa* (wild type), in which the promoter, *lac* polymerase is induced by the addition of IPTG.

An overnight starter culture of 1 colony in 10 ml LB (plus Carbenecillin 500 mg/l) grown aerobically at 37 °C, was used to inoculate 1L LB medium and grown at 37 °C plus Carbenecillin 500 mg/l at the 37 °C, until the OD<sub>578</sub> reached approximately 0.9. IPTG was then added to a final concentration of 1 mM. Cultures were further incubated as appropriate at 37 °C or 16 °C, before harvesting by centrifugation (10 000 x g, 4 °C). The bacterial pellets were resuspended in the appropriate binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM buffer). Proteins were then purified in the same way, using metal chelate chromatography. The proteins were overproduced aerobically (shaking) and anaerobically (no shaking) covering the culture with sterile mineral oil and the addition of sterile sodium nitrate at a final concentration of 25 mM.

### 2.3.3.3 Sonication of bacteria

Prior to purification, the crude cell fractions were sonicated using a SonicsVibracell Ultrasonic processor, with an output between 20 and 30 watts. To the pellet, 1 complete, mini, EDTA-free protease inhibitor cocktail tablet was added per litre to be sonicated to prevent protein degradation. These inhibit serine and cysteine proteases, and the absence of EDTA allows for the stability and function of metal dependant proteins. Sonication was performed in burst of 30 secs on and 30 sec off, approximately 6 times, or until the cells were sufficiently lysed. The lysed cells were then centrifuged at 16,000 rpm for 20 min at 4 °C. The resultant supernatants containing soluble protein were then subject to purification by metal chelate chromatography.

### 2.3.3.4 Purification by Metal chelate chromatography

The overproduced recombinant protein is fused to an N-terminal Histag. This allows us to purify the protein by exploiting the affinity that histidine has for chelating to metals such as  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Fe}^{3+}$ , and  $\text{Ni}^{2+}$ . Therefore, Chelating Sepharose fast flow resin, with immobilised  $\text{Ni}^{2+}$  will selectively bind and retain the recombinant protein whilst unbound proteins are washed away. The bound recombinant proteins are then isolated from the column by competition elution with imidazole. For a detailed methodology please refer to the Amersham Biosciences Chelating Sepharose Fast Flow instruction manual.

Purification of the various proteins has been detailed individually in the appropriate chapters for brevity. To summarise columns containing ~5mls of resin were charged using 0.1M  $\text{NiSO}_4$  (15 ml) and equilibrated with the appropriate binding buffer for the protein (15 ml). The supernatant from the sonicated extract (Section 2.4.3.3) was then applied to the column and allowed to flow through. The columns were washed with binding buffer, and wash buffers of increasing imidazole concentration (50 mM then 100 mM). The proteins were eluted in 2 ml fractions using elution buffer (400 mM imidazole), and stored on ice. The presence of protein in the fractions was detected by the BioRad protein assay (Section 2.3.1) and SDS-PAGE (Section 2.3.4).

### **2.3.3.5 Histag Cleavage**

One unit of thrombin is required to cleave one milligram of protein. Therefore. The concentration of protein was estimated by  $A_{280}$ , and an appropriate volume of thrombin was added to the protein solution and left overnight at 4 °C. Cleavage was checked by running both samples on a SDS-PAGE gel.

### **2.3.3.6 Buffer Exchange / Desalting**

Purified protein was desalted or changed into a different buffer using pre-packed Sephadex G25, pD10 columns, details can be found in the manual (Amersham). The column is pre-equilibrated with 50 ml of the buffer into which the protein is to be exchanged, before adding 2.5 ml of the protein sample. The protein is then eluted into 3.5 ml of the new buffer

### **2.3.3.7 Gel filtration / Fast protein liquid chromatography**

Where possible, to further remove protein aggregates and to estimate native molecular weight of the protein, a sufficiently concentrated solution of protein was loaded onto pre-packed high flow, high resolution Superdex 75 column (molecular mass range  $5 \times 10^2$  to  $7 \times 10^4$  or) or High Prep Sephadex 200 column, in conjunction with Äkta FPLC (all from Amersham). These were pre equilibrated with 1.5 - 2 column volumes of column buffer as per the protein. The protein is then eluted at 0.5 ml/min or 1.0 ml/min into 0.5 ml or 1 ml fractions with buffer as appropriate to the protein. Fractions were analysed by SDS-PAGE and the purest fractions pooled. The columns are then washed with wash buffer and stored in storage buffer. Details of all the buffers used can be found in Section 2.1.6.4.

### **2.3.4 Poly- Acrylamide Gel Electrophoresis (PAGE) of proteins**

The composition of all the buffers, solutions and stacking and running gels are detailed in Section 2.1.6.2.

#### **2.3.4.1 SDS-PAGE of proteins**

In order to visualise the purity and integrity of the resultant overproduced proteins, SDS-PAGE was performed according to the method outlined by Laemmli (Laemmli 1970). This is widely used for the separation of proteins by molecular weight and charge. Gels typically consist of a resolving gel (pH 8.8) for protein separation and a stacking gel (pH 6.8) that focuses the proteins prior to entering the resolving gel. The final percentage of acrylamide was determined according to the molecular weight of the protein to be identified. Since all the proteins in this study were in the region of 19 to 46 kDa, a 12.5 % acrylamide gel was sufficient for resolution.

Protein samples were denatured by addition of an equal volume of SDS buffer to the sample and boiling for 5 minutes. Once the gel had polymerised, the gel plate sandwich was placed into an electrophoresis chamber (Atto Dual Mini Slab AE6450) flooded with SDS running buffer. Samples were loaded into the wells alongside a Dalton VII molecular weight marker, with known protein sizes as a standard. Electrophoresis was performed at a constant voltage of 200 V until the loading dye front (bromophenol blue) reached the end of the gel plate. The gels were removed carefully from the apparatus and stained with coomassie blue stain and de-stained using water. They are then dried between two sheets of cellophane (BioRad), for convenient storage.

#### **2.3.4.2 Native-PAGE of proteins**

Native gels or non-denaturing gels were prepared as described in Section 2.1.6.2, with the exception of the non-denaturing nature of the loading sample buffer (no SDS) and the samples not being boiled. Thus in this type of separation we may anticipate the presence of multimeric unfolded states native to the protein. Compact conformations have a higher mobility through the gel whereas larger structures, like oligomers, have lower mobility. The samples are directly loaded onto the gel after adding the loading sample buffer. A native molecular weight marker is used as a standard. Electrophoresis is performed on the same apparatus at a constant voltage of 100 V. After electrophoresis, the gels are removed carefully from the apparatus

and stained with coomassie blue stain and de-stained using water. They are then dried between two sheets of cellophane (BioRad), for convenient storage.

### **2.3.5 Western blotting of Histagged proteins**

The composition of all the buffers, solutions and stacking and running gels are detailed in Section 2.1.6.3.

#### **2.3.5.1 SDS-PAGE for Western blotting of proteins**

Gels and samples were prepared and run as in the same way as before as detailed in Section 2.3.4.1, except the electrophoresis was performed at a constant voltage of 150 V and a pre stained high or low molecular weight marker was used instead of Dalton VII as a standard.

#### **2.3.5.2 Western blotting**

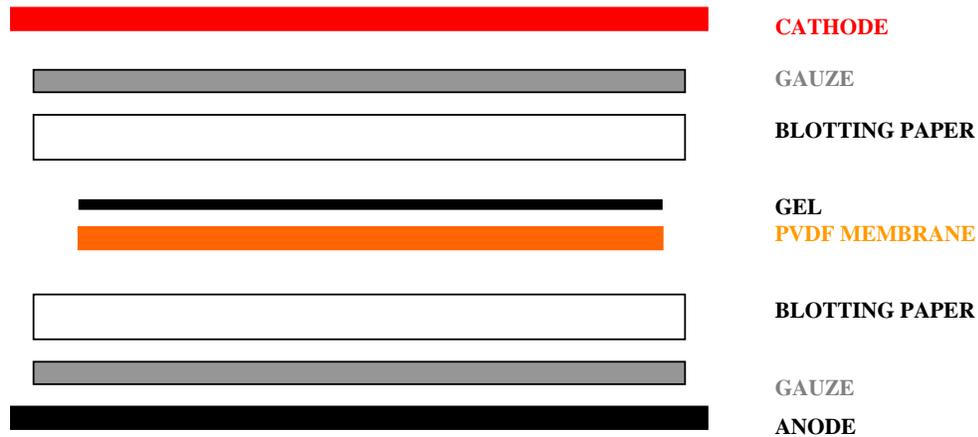
Proteins separated on a denaturing polyacrylamide gel were electrophoretically transferred onto a Hybond-P, hydrophobic polyvinylidene difluoride (PVDF) membrane (Hybond-P (Pharmacia)) membrane using the method outlined by (Towbin, Staehelin et al. 1979).

The unstained SDS-PAGE gels, blotting paper, fibre pads were all pre-soaked in cold transfer buffer (20 % Methanol, 192 mM Glycine, 25 mM Tris). The PVDF transfer membrane (Hybond-P (Pharmacia)) was saturated in 100 % methanol, then water before soaking in transfer buffer.

The blot was arranged as shown in Figure 2.5 ensuring no air bubbles were present, and placed in a BioRad gel cassette with an ice block, in ice-cold transfer buffer. The proteins were transferred to the membrane by passing 100 V for 1 h in to the cassette. After the transfer, the PVDF membranes were blocked for 2 hrs in 5 % non-fat milk in PBS (or overnight at 4 °C) to block non-specific sites, on a platform shaker (140 mM NaCl, 3 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 mM KH<sub>2</sub>PO<sub>4</sub>).

**Figure 2.6: Assembly of the western blotting apparatus**

The Western blot must be assembled in the order here, starting from the black anode end assembling upwards. The membrane is closest to the anode, then the gel, ensuring there are no air bubbles and that the sandwich is placed in the tank facing the right way. The proteins are then transferred to the membrane at 100V for 1hr.

**2.3.5.3 Probing for Histagged proteins**

Once the membrane was been blocked for non specific binding in milk-PBS-solution, (2 hrs or overnight) to decrease any background activity, the membrane was probed for the primary antibody, left overnight (or for 2 hrs) at 4 ° C in 5 % non-fat milk in PBS, containing 1/1000 primary antibody (mouse anti-his monoclonal antibody, Novagen), with constant gentle agitation.

After washing the membrane with PBS and three times with Tris-NaCl wash buffer (50 mM Tris-HCl pH7.5, 150 mM NaCl), the membrane was incubated with the secondary antibody, goat anti-mouse IgG alkaline phosphatase conjugate, for 1 hr at room temperature at a working dilution of 1:5,000 in 5 % non-fat milk in 50 mM Tris, 150 mM NaCl.

The membrane was again washed with Tris-NaCl buffer thoroughly before proteins were visualised by incubating the membrane with one tablet of Sigma Fast BCIP/NBT (5-Brom-4-choro-3-indolyl phosphate / nitroblue tetrazolium), dissolved in 10 ml of water, a substrate for alkaline phosphate. The colour, a dense blue, was allowed to develop and then stopped by washing in water. They were then allowed to dry and stored in away from light.

## 2.4 Isolation and production of tetrapyrroles

### 2.4.1 UltraViolet-Visible Spectroscopy

All UV-Vis spectra were recorded on a Varian Cary 50 Bio UV-Vis spectrophotometer over a range of 300 – 650 nm.

### 2.4.2 Production of precorrin-2 and sirohydrochlorin *in situ* from cell extracts

*In situ* production of precorrin-2 and sirohydrochlorin was performed under a nitrogen atmosphere, in a glove box (Belle Technology) with less than 2 ppm oxygen. This can be performed by synthesising from either ALA, or PBG available commercially (Sigma).

This was achieved by incubating the crude soluble extract of recombinant *E. coli* strains over-producing the enzymes required for the transformation of ALA into PC-2 or SHC (donated by Dr Helen Leech). The reaction mixture was incubated as follows: 10.0 mg ALA in a total volume of 40 ml of 50 mM Tris-HCl buffer, pH 8.0, containing 5 mg of ALA-dehydratase (HemB, *B. megatarium*), 5 mg of purified PBG deaminase (HemC, *B. megatarium*), 1 mg purified uroporphyrinogen III synthase (HemD, *B. megatarium*), 5 mg of purified UIII methyltransferase (CobA, *Ps. denitrificans*) and 10 mg of SAM. If proceeding from PBG, this is done by incubating 2.5 mg PBG in a total volume of 40 ml of 50 mM Tris-HCl buffer, pH 8.0, containing 5 mg of purified PBG deaminase (HemC, *B. megatarium*), 1 mg purified uroporphyrinogen III synthase (HemD, *B. megatarium*), 5 mg of purified UIII methyltransferase (CobA) and 10 mg of SAM. All powdered substances e.g. SAM, ALA, or PBG were suspended in 1 ml 50 mM Tris-HCl buffer, pH 8.0, just prior to addition and added to a total volume of 40 ml. The reaction mixture is then freeze thawed with liquid nitrogen, under vacuum to eliminate any air and left overnight under a strictly anaerobic environment at 22 °C to allow it to reach completion. Sirohydrochlorin was generated using the same enzyme cocktail as described above, except that 5 mg of PC-2 dehydrogenase (SirC, *B. megatarium*) and 13 mg of NAD<sup>+</sup> were added to the incubation.

Prior to use, the mixture is filtered (2 µm) and the UV-Vis spectrum recorded to estimate PC-2 or SHC concentration.

### 2.4.3 Production of Precorrin-2 and sirohydrochlorin from purified proteins

This method of production of PC-2 and SHC gives a much higher yield and purer form of the substrate. For the production of PC-2, pETcocoABCD (*hemB*, *hemC*, *hemD*, *cobA*) is used and for SHC, pETcocoABCDC (*hemB*, *hemC*, *hemD*, *cobA*, *sirC*) is used.

Essentially both methods are identical except for the addition of  $\text{NAD}^+$  to the reaction mixture. Each plasmid is transformed into BL21\* DE3 pLysS and plated on to LB-agar plates with ampicillin and 0.2 % glucose. A colony is used to inoculate a 10 ml LB starter culture, which is then used to inoculate 1L of LB and grown at 37 °C until an OD of 0.2 at which point a final concentration of 0.02 % Arabinose was added to the culture. This was further incubated at 37 °C until an OD of 0.6 (2 hrs) when the culture was induced with a final concentration of 0.4mM IPTG and left overnight at 16 °C. Cells were harvested by centrifugation and resuspended in 10 ml 50 mM Tris-HCl buffer, pH 8.0. This was divided into four 2.5 ml aliquots, each being sufficient for the purpose and stored at -20 °C.

Pellets were defrosted as required, sonicated, centrifuged and the soluble cell extract applied to a metal chelate (Ni) column. Proteins were eluted at 400 mM imidazole into 1ml fractions and pooled. Purified protein was taken into the anaerobic glove box and passed through a pD10 column to remove any oxygen. 10 mg ALA, 5 mg SAM and 10 mg  $\text{NAD}^+$  were taken into the glove box in powder form. The ALA and SAM were each dissolved in 50 mM Tris HCl pH 8.0, 0.1M NaCl to 500  $\mu\text{l}$  and combined to give a final total volume of 1 ml and the pH adjusted to pH 8 using KOH (10  $\mu\text{l}$ ). The  $\text{NAD}^+$  was dissolved in 50 mM Tris HCl pH 8.0, 0.1M NaCl to 1 ml. Once the proteins were passed through a pD10 column they were used quickly to make the following reaction mixtures:

<u>For PC-2</u>	<u>For SHC</u>
3 ml buffer 50 mM Tris HCl pH 8.0, 0.1M NaCl	
1 ml ABCD	1 ml ABCDC
1 ml ALA (10 mg) + SAM (5 mg) in 50 mM Tris HCl, 0.1M NaCl (above)	
	1ml $\text{NAD}^+$ at 10 mg/ml

The mixture is left on a warmed heat block (turned off) overnight. There is no need to filter the mixture prior to use but the UV-Vis spectrum is used to estimate PC-2 or SHC concentration.

#### **2.4.4 Production of haem $d_1$ and intermediates**

Two methods were employed in the production of possible intermediates in the haem  $d_1$  pathway. Primarily, *in vivo* accumulation was attempted with the expression of multigene plasmids constructed in Table 2.4, both within *E. coli* and *Ps. aeruginosa*. Secondly, these same plasmids were expressed and reacted to pre-made PC-2 and SHC under anaerobic conditions.

##### **2.4.4.1 *In vivo* accumulation of intermediates**

Plasmids containing *nir* gene combinations of interest were transformed into *E. coli* and *Ps. aeruginosa* as appropriate and subsequently used to inoculate 1 L of LB and grown at 37 °C for 6-8 hrs until the OD reached 0.6. At this point, 0.4 mM IPTG in *E. coli* and 1 mM IPTG in *Ps. Aeruginosa* was added, as well as the addition of ALA (10 mg/ml) and SAM (10 mg/ml). The cultures were then incubated overnight and harvested by centrifugation. The media was kept as this stage and the pellets were resuspended in 50 mM Tris HCl pH 8.0, 0.1M NaCl and sonicated for 30 sec on, 30 sec off three times at an amplitude of 10  $\mu$ m. The resultant soluble cell extract and media was applied to a DEAE sephacel column or RP18 column aerobically or anaerobically. The columns were then washed as outlined for the type of column in Section 2.4.6 and Section 2.4.7, and any accumulated compounds were analysed by a UV-Vis spectrophotometer (HP-352 photodiode array).

##### **2.4.4.2 *In vitro* accumulation of intermediates**

PC-2 and SHC were synthesised anaerobically as described in Section 2.4.2 and Section 2.4.3. Accordingly the *nir* genes singularly or combinations thereof were expressed in either *E. coli* or *Ps. aeruginosa*. The single expressed genes were purified if they were tagged, in a buffer as appropriate to the stability of the protein. Alternatively if they were not tagged or the multigene plasmids were being expressed, the pellet was resuspended in 50 mM Tris HCl pH 8.0, 0.1M NaCl and the soluble fraction used after sonication. The purified protein or soluble cell extract was taken into the anaerobic glove box and passed through a pD10 column to remove any oxygen from the preparation. Accordingly, each condition as described in Chapter 5 was incubated with PC-2 or SHC overnight at 30 °C with or without potential cofactors as detailed below. The resultant mixture was applied to a DEAE sephacel column or RP18 column, which had been pre-equilibrated the same buffer in an anaerobic glove box. The columns were then washed as outlined for the type of

column in Section 2.4.6 and Section 2.4.7, and any accumulated compounds were analysed by a UV-Vis spectrophotometer (HP-352 photodiode array).

10 mg/ml NAD<sup>+</sup>  
10 mg/ml NADP  
10 mg/ml NADPH  
10 mg/ml SAM  
1 mM Ferrous (FeII) Sulphate  
1 mM Sodium Dithionate

Each potential cofactor was taken into the anaerobic glove box in powder form and dissolved in 50 mM Tris-HCl pH 8.0, 0.1M NaCl to 6 ml.

#### **2.4.5 *In vitro* sirohydrochlorin coupled SUMT assay**

The assay was performed using purified SUMT, and a cocktail of HemB, HemC, HemD, and SirC. The SUMT (NirE or CobA) was overproduced and purified by metal chelate chromatography separately. A plasmid containing all four genes HemB, HemC, HemD, and SirC, pET14b-BCDC, was similarly overexpressed and the resultant overproduced recombinant proteins were purified in the same fashion and eluted at 400 mM imidazole. These were both passed through a pD10 column in the anaerobic glove box to remove any oxygen changing the buffer to 50 mM Tris HCl pH 8.0, 0.1 M NaCl. 20 mg ALA, 20 mg SAM and 26 mg NAD<sup>+</sup> were taken into the glove box in powder form and each dissolved in 50 mM Tris-HCl pH 8.0, 0.1M NaCl to 1 ml.

All Proteins and cofactors were added in excess, except for the SUMT, to a total volume of 1 ml in a cuvette in 50 mM Tris-HCl buffer, 100 mM NaCl, pH 8.0, buffer. This allowed the reaction to be followed anaerobically by UV-Vis spectroscopy. A total of 20 µg ALA, 20 µg SAM, 26 µg NAD<sup>+</sup>, 200 µg BDCD was added to the reaction mixture. The SUMT was added last, incubated at room temperature and the reaction was followed at 5-minute intervals over a total of 60 minutes. At each interval a UV-Vis scan was taken and the absorbance at 376 nm was followed to monitor the generation of SHC since SHC has an absorption maxima at 376 nm. The concentration of SUMT and the presence of the cofactors was varied in the reaction and monitored by UV-Vis spectroscopy. The specific

activity was calculated using the extinction coefficient of SHC as  $2.4 \times 10^5 \text{ M}^{-1}\text{cm}^{-1}$ , using the equation below:

$$\text{Specific activity (nmoles/min/mg)} = \frac{\text{initial rate}}{\epsilon\text{SHC} \times \text{enz mass (g)}} \times 10^9$$

$$\text{Initial rate (abs/min)} = \frac{\Delta \text{ abs}}{\Delta \text{Time (min)}}$$

#### **2.4.6 Reverse Phase chromatography separation of pigments**

The resultant soluble cell extract was subject to reverse-phase chromatography that uses a non-polar stationary phase. Previously most chromatography was done with hydrophilic surface chemistry and had a stronger affinity for polar compounds but here polar compounds are eluted first while non-polar compounds are retained - hence “reversed phase”

Approximately 3 g of RP-18 (particle size 5  $\mu\text{m}$ ) is charged with 10ml hexane, enough to make a 2 cm column, allowing it to be poured into a column. It is then washed with one column volume of 95 % analar ethanol, one column volume of ddH<sub>2</sub>O, and then soluble cell extract is applied to the column. Half a column volume of ddH<sub>2</sub>O is applied and the flowthrough collected. Any pigments or accumulated compounds were eluted at 25 % ethanol and collected and analysed by a UV-Vis spectrophotometer (HP-352 photodiode array).

#### **2.4.7 Ion exchange chromatography separation of pigments – DEAE Sephacyl**

DEAE Sephacyl is a weak anion exchanger based on beaded cellulose. The ion exchange group is diethylaminoethyl, which remains charged and maintains consistently high capacity over the entire working range, pH 2–9.

The resultant soluble cell extract was applied to a DEAE sephacel column pre-equilibrated the same buffer in an anaerobic glove box. The columns were then washed with two column volumes of the same buffer and any accumulated compounds were eluted, by application of 50 mM Tris HCl pH8.0 with 2 M NaCl and analysed by a UV-Vis spectrophotometer (HP-352 photodiode array).

## **2.5 Crystallography - Hanging drop method**

Crystallisation trials were carried out using the hanging drop vapour diffusion method (McPherson 1976) . Recombinant protein was purified on a metal chelate (Ni) column and by FPLC on a size exclusion column as detailed previously. The protein was further concentrated using vivaspin columns (Vivascience). Each reservoir of a 24-well crystallisation plate was filled with 500 µl of the buffers from Hampton Research Crystal Screen I<sup>TM</sup> and Crystal Screen II<sup>TM</sup>. 2 µl of concentrated protein samples were mixed with 2 µl of reservoir buffer on a cover slip. The cover slip was inverted, placed over the well of the same buffer that was sealed with the silicone grease. All the crystallisation trials were carried out at 18 °C or s appropriate in a controlled temperature room.

**Characterisation of NirE;  
an *S*-Adenosyl-L-methionine  
Uroporphyrinogen III Methyltransferase**

### 3. Characterisation of NirE – a *S*-Adenosyl-L-methionine uroporphyrinogen methyltransferase (SUMT)

#### 3.1. Introduction

Haem  $d_1$  belongs to a family of tetrapyrroles that include vitamin B<sub>12</sub>, haem, chlorophyll and sirohaem, all of which branch from the first macrocyclic intermediate uroporphyrinogen III (UIII) (Chapter 1, Figure 1.2). From this point the pathway diverges to form the various tetrapyrroles, where decarboxylation of its acetate groups directs flux toward haem and chlorophyll synthesis, whereas methylation of the UIII macrocycle initiates synthesis toward sirohaem, vitamin B<sub>12</sub>, and haem  $d_1$ . The UIII macrocycle is shown in Chapter 1, Figure 1.1.

*S*-Adenosyl-L-methionine uroporphyrinogen III methyltransferase (SUMT) defines a group of enzymes that catalyse the methylation of the UIII to produce precorrin-2 (PC-2, also known as dihydrosirohydrochlorin). Sirohaem, vitamin B<sub>12</sub>, coenzyme F<sub>430</sub> and haem  $d_1$  all require this reaction as part of their synthesis and so are derived from PC-2. Thus, SUMT enzymes are also positioned at a key branch point in the synthesis of modified tetrapyrroles directing flux towards these molecules and portray some interesting regulatory features.

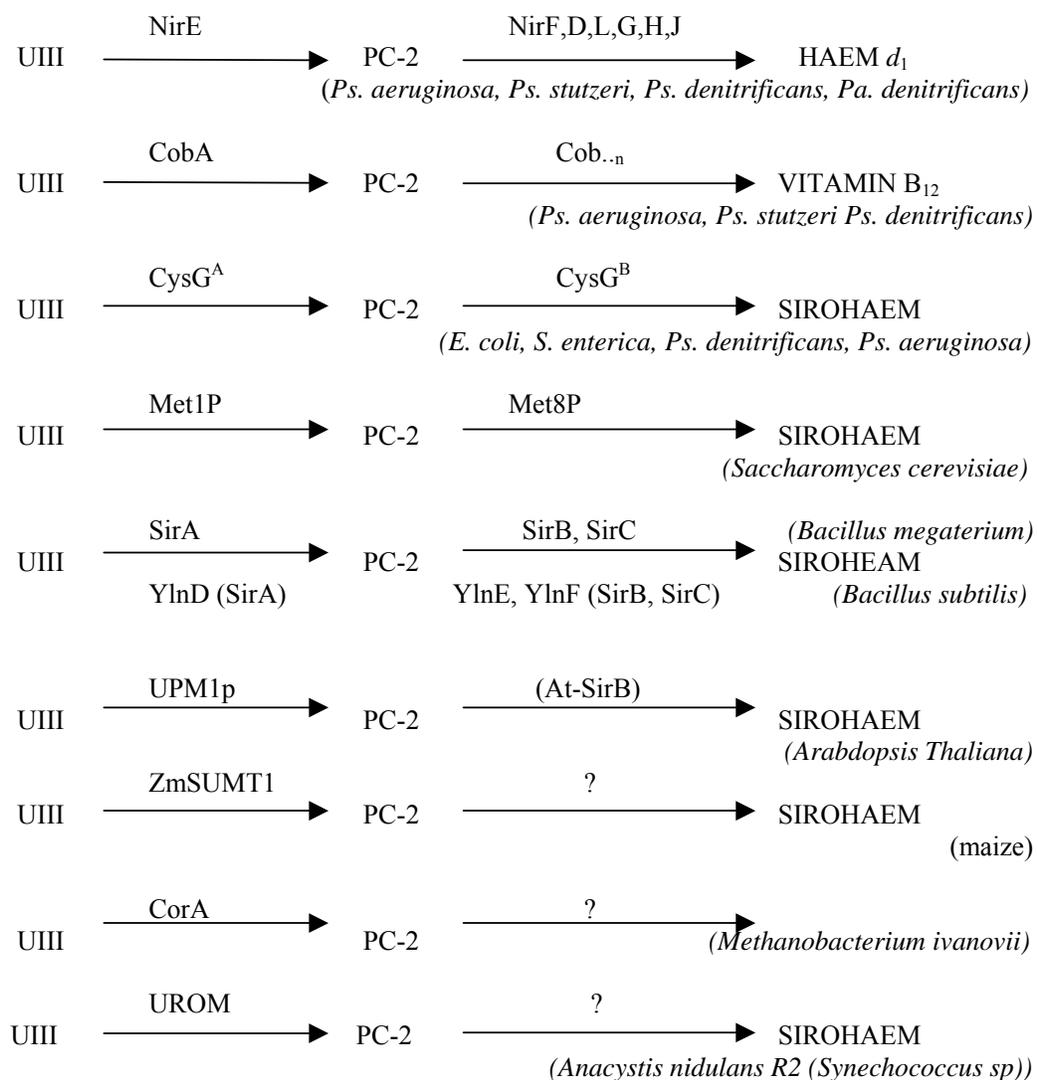
All SUMT enzymes utilise *S*-adenosyl-L-methionine (SAM) as the methyl donor to insert two methyl groups onto the UIII macrocycle to yield PC-2, by the subsequent methylation of UIII at the base of the acetate side chains at positions C2 and C7 to produce PC-2. This reaction proceeds via a mono-methylated intermediate precorrin-1 (PC-1) with the concomitant release of the product *S*-adenosyl-L-homocysteine (SAH). Precorrin-1 is in turn methylated to produce the *bis*-methylated PC-2 (Figure 1.19) (Warren, Stolowich et al. 1990).

There exist a plethora of SUMTs that have been discovered in different organisms and plants that perform this transformation; these are summarised in Figure 3.1. The most well known and well studied is CobA, which is associated with cobalamin biosynthesis. This was the first SUMT to be characterised from *Ps. denitrificans* in 1989 (Blanche, Debussche et al. 1989). Since then orthologues have been identified in *B. megatarium* as SirA (Raux, Leech et al. 2003), *E. coli* as CysG<sup>A</sup> (Warren, Roessner et al. 1990), CorA in *M. ivanovii* (Blanche, Robin et al. 1991) and as

UROM *Anacystis nidulans* R2 (also called *Synechococcus* sp. PCC7942) (Jones, Jenkins et al. 1994) and Met1P in *S. cerevisiae* (Raux, McVeigh et al. 1999). SUMTs also exist in plants as UPM1p in *Arabidopsis Thaliana* (Leustek, Smith et al. 1997) and ZmSUMT1 in maize (Sakakibara, Takei et al. 1996). NirE is also postulated to be part of the SUMT family. It has been found in *Ps. aeruginosa* (Kawasaki, Arai et al. 1997) and in *Pa. denitrificans* (Chang and Wu 1986). At present, there exists little information on NirE and this chapter reveals new findings.

**Figure 3.1: SUMTs and their associated downstream enzymes – a diagrammatic representation.**

Outline of the various organism and plant SUMTs transforming UIII to PC-2. The majority are involved in sirohaem synthesis. *B. subtilis* YlnD, YlnE, YlnF are renamed SirA, SirB, and SirC respectively to reflect their role in sirohaem synthesis.



CobA from *Ps. denitrificans* is responsible for the C2 and C7 corrin ring methylations in the conversion of UIII into PC-2 in the aerobic and anaerobic pathways to cobalamin of some thirty enzymes (Blanche, Debussche et al. 1989). Sirohaem synthesis is much less complicated than cobalamin requiring only CysG, a unique multifunctional enzyme encompassing all the properties to transform UIII into sirohaem, in both *E. coli* and *S. enterica* (Peakman, Crouzet et al. 1990; Warren, Roessner et al. 1990; Wu, Siegel et al. 1991). *E. coli* and *S. enterica* CysG have 90 % homology to each other and have two functional domains; CysG<sup>B</sup> performing dehydrogenation and ferrochelatase activity and CysG<sup>A</sup> which mediates SUMT activity (Warren, Bolt et al. 1994).

As opposed to the fused multifunctional CysG, *B. subtilis* and *B. megaterium* use three proteins for the conversion of UIII into sirohaem present in a distinct gene cluster. SirA is confirmed to confer SUMT activity; SirC is a PC-2 dehydrogenase and SirB a ferrochelatase completing the sirohaem pathway (Johansson and Hederstedt 1999; Leech, Raux-Deery et al. 2002; Raux, Leech et al. 2003).

The enzymes identified for the transformation of UIII into sirohaem in *S. cerevisiae* are Met1p and Met8p. (Hansen, Muldbjerg et al. 1997; Raux, McVeigh et al. 1999). Met1p is similar to CysG<sup>A</sup> conferring SUMT activity and Met8p is a bifunctional enzyme with NAD<sup>+</sup> dependent PC-2 dehydrogenase and ferrochelatase activity. Thus, the roles of Met1p and Met8p are considered to be isofunctional with CysG (Raux, McVeigh et al. 1999). MET1p from *S. cerevisiae* and SirA from *Bacilli*, which function independently, are homologous to CysG<sup>A</sup> (Warren, Bolt et al. 1994; Raux, McVeigh et al. 1999). MET8p and SirC also display similarity to CysG<sup>B</sup>, but SirB has no sequence identity with CysG<sup>B</sup>, suggesting CysG is somewhat unique in that the gene fusion event that created this trifunctional enzyme (Warren, Roessner et al. 1990; Warren, Bolt et al. 1994). Although *E. coli* does not produce vitamin B<sub>12</sub>, *S. enterica*, does make vitamin B<sub>12</sub> anaerobically, suggesting CysG in *S. enterica* is common to both pathways (Raux, McVeigh et al. 1999).

UPM1p from *A. thaliana* is identified as an SUMT as part of the sirohaem biosynthetic pathway in plants and shown to be ~40kDa with ~43% homology to *E. coli* CysG<sup>A</sup>. The search for a sirohydrochlorin dehydrogenase continues as it is the only missing component of the sirohaem pathway in *A. thaliana* (Leustek, Smith et al. 1997; Raux-Deery, Leech et al. 2005). ZmSUMT1 was isolated and identified to

be a SUMT. Its characterisation revealed it to be homodimeric with a molecular mass of 52 kDa by gel-filtration chromatography. It displays about 50% similarity to the *E. coli* CysG<sup>A</sup> SUMT. (Sakakibara, Takei et al. 1996; Fan, Wang et al. 2006).

Resolving the topology of CobA from *Ps. Denitrificans* (Vevodova, Graham et al. 2004), has given great insight in to the structure function relationship of SUMT activity. It is a homodimer with a subunit mass of 29 kDa and a low catalytic efficiency (38/h) (Blanche, Debussche et al. 1989). The position of the SAH binding site in CobA is the same in CysG from *S. enterica* (Stroupe, Leech et al. 2003), also a homodimer. A total of nineteen conserved residues have been attributed to SAH/SAM binding, including the glycine rich nucleotide binding region denoted by GXGXG (GAGPGD) (Vevodova, Graham et al. 2004). The structure of the CysG<sup>A</sup> domain resembles the structure of CbiF, also a homodimer, a cobalt-precorrin-4 methyltransferase enzyme from *B. megaterium* in the anaerobic cobalamin pathway (Schubert, Wilson et al. 1998), and is concomitant with other cobalamin biosynthetic methyltransferases (Frank, Deery et al. 2007).

Substrate inhibition of CobA from *B. megaterium* and *Ps. denitrificans* is observed at UIII concentrations above 2  $\mu\text{M}$ , values below those found intracellularly indicating a regulatory role in SUMT activity and flux through the cobalamin pathway by way of feedback inhibition. CobA was also found to be sensitive to S-adenosyl-L-homocysteine (SAH) but not to any other vitamin B<sub>12</sub> intermediates, providing another regulatory mechanism. (Blanche, Debussche et al. 1989; Crouzet, Cauchois et al. 1990). *B. megatarium* SirA was observed to have substrate inhibition with UIII concentrations ([UIII]) above 0.5  $\mu\text{M}$ . However, it was desirable to obtain an SUMT that does not exhibit substrate inhibition to allow greater flux through these pathways. Such a SUMT was found in a methanogen called *M. ivanovii* which did not display substrate inhibition until a [UIII] close to 20  $\mu\text{M}$  was reached (Blanche, Robin et al. 1991). It is perceived that since methanogens require such high levels of coenzyme F<sub>430</sub> that inhibition does not occur at low UIII substrate concentrations. Therefore in comparison, SUMT substrate inhibition as observed in *B. megaterium* and *Ps. denitrificans* can be attributed to the lower requirement of cobalamin in these organisms (Robin, Blanche et al. 1991; Fan, Wang et al. 2006).

The *nirE* gene was first discovered in *Pa. denitrificans*, downstream of *nirS* (de Boer, Reijnders et al. 1994). It has since been found in *Ps. aeruginosa* (Kawasaki, Arai et al. 1997) and in *Ps. stutzeri* (Palmedo, Seither et al. 1995). Vollack *et al* 1998, used Southern hybridisation to probe for denitrification genes in *Ps. aeruginosa* and found two signals for *nirE*. Further analysis revealed that one persisted more and was part of the *nir* gene cluster (Vollack, Xie et al. 1998). The signal not associated with the *nir* cluster is presumably another SUMT, CobA or CysG. This demonstrates the importance of enzyme multiplicity as a factor for tight regulation of flux towards the biosynthesis of different tetrapyrroles in *Ps. aeruginosa*.

NirE has been reported to be a SUMT of the haem  $d_1$  pathway on the basis of sequence similarity to well the characterised SUMTs discussed previously. However the activity of NirE has never been demonstrated to transform UIII to PC-2 experimentally. Only from mutational and complementation studies, has it been shown that *nirE* is essential for haem  $d_1$  biosynthesis. Palmedo *et al* used mutational analysis in *Ps. stutzeri* to show that if any one or all *nir* genes are defective it results in an inactive semi apofrom of cytochrome  $cd_1$ , lacking haem  $d_1$  but containing haem  $c$  (Palmedo, Seither et al. 1995). Kawasaki *et al* showed the same in *Ps. aeruginosa* by mutational analysis as well as complementation, both indicating *nirE* is essential for haem  $d_1$  biosynthesis (Kawasaki, Arai et al. 1997).

The connection of haem  $d_1$  to the tetrapyrrole biosynthetic pathway was established by the unique stereo-specificity and origins of the macrocycle and its peripheral groups. The source of the methyl groups at C2 and C7 is SAM and implicates C-methylation as part of the route to haem  $d_1$  (Yap-Bondoc, Bondoc et al. 1990). Experiments using labelled ALA showed the haem  $d_1$  macrocycle is derived from ALA, including the carbonyls at C3 and C8 except for the methyls at C2 and C7 (Matthews 1993). The exact stereo-specificity of the methyl groups on haem  $d_1$  was found to be exactly the same as vitamin B<sub>12</sub>, sirohaem and its intermediates, in the absolute configuration of 2*R*, 7*R*. This further implicates that methylation occurs as part of the haem  $d_1$  pathway consistent with the transformation of UIII to PC-2 (Micklefield 1997).

Thus, methylation of UIII forming PC-2 is common to all these pathways differing in the enzyme utilised within the organism as well as the pathway. Indeed there exists an evolutionary advantage to have more than one SUMT in a single organism, as increased levels of substrate, and different levels of substrate inhibition for each SUMT allows tight regulation of tetrapyrrole biosynthesis. The presence of NirE may pose some advantage or specificity in its activity and is as yet to be discovered.

In this chapter, *nirE* is cloned, overexpressed and the resultant recombinant protein purified and characterised from two different systems; the first overproduces NirE heterologously in *E. coli* and the second uses *Ps. aeruginosa* for the overproduction of NirE homologously. The roles of a number of residues are discussed with respect to their conservation across the family of SUMTs. NirE is shown for the first time to be a functionally active enzyme, performing the transformation from UIII to PC-2, both *in vivo* and *in vitro* as part of the haem  $d_1$  pathway.

## RESULTS AND DISCUSSION

### 3.2. Sequence identity of NirE to various SUMTs

Sequence identity of NirE from *Pseudomonads* to CobA of *Ps. denitrificans* is documented in the region of 45.1 % with 56.8 % identity to CysG from *E. coli* and 57.6 % to CysG from *S. enterica* (Palmedo, Seither et al. 1995; Kawasaki, Arai et al. 1997).

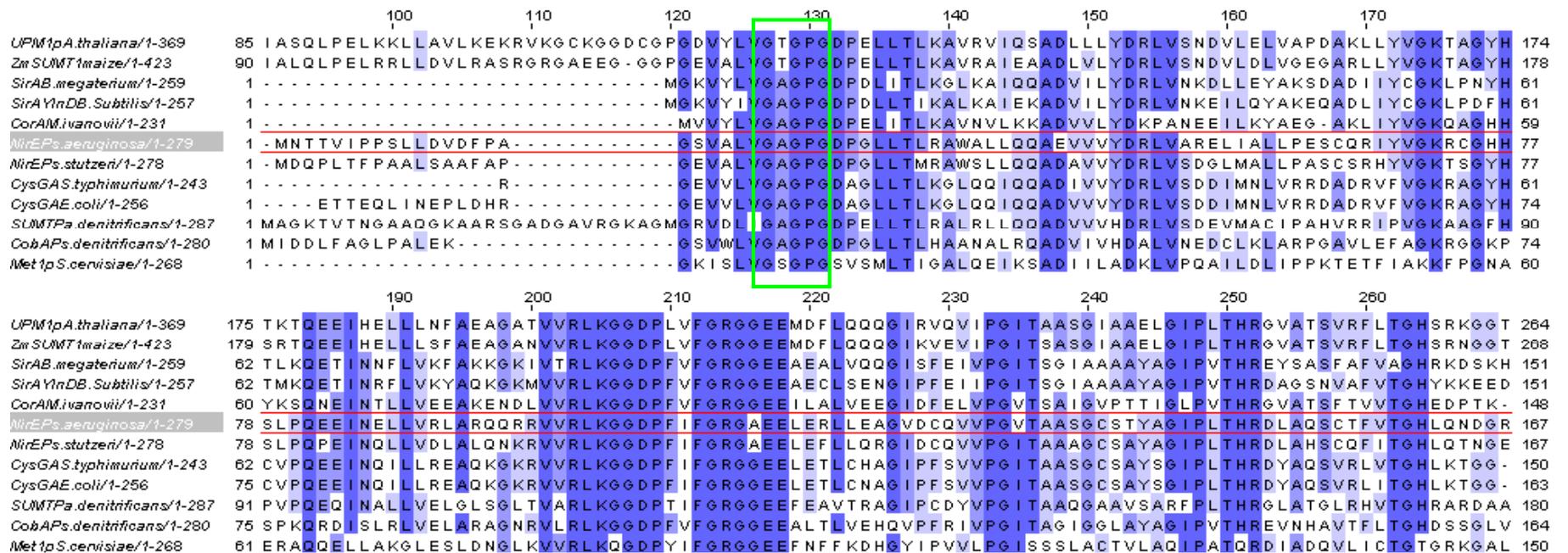
Since CobA is well characterised, NirE has been aligned with CobA showing a similarity of 50.3 % and an amino acid identity of 37.3 %, shown in Figure 3.2. The aligned nucleotide sequences of CobA from *Ps. denitrificans* (843 base pairs) with NirE from *Ps. aeruginosa* (839 base pairs) an identity of 56.0 %. The similarity of NirE to the well-characterised SUMT enzymes is shown in a multiple alignment (Figure 3.3). One can see there are 31 completely conserved amino acid residues across the various SUMTs. A highly conserved SAM binding site is evident in NirE as with the other SUMTs with a GXGXG region, highlighted by the green box in Figure 3.3. The sequence identity of these SUMT enzymes to NirE and relative to each other has been summarised in Table 4.1, on the bases of pair wise alignments using ClustalW. Clearly they have a high degree of similarity and identity to each other portraying an average sequence identity of 45 %. The lowest with Met1p at 30 % and the highest being with NirE from *Ps. stutzeri* at 63 %.

The alignment of NirE with CysG has been performed with the CysG<sup>A</sup> C-terminal domain, as CysG<sup>B</sup> shows no sequence similarity to NirE (or MET8p-data not shown). This predicts NirE has a sole function as a *S*-adenosyl-L-methionine uroporphyrinogen UIII methyltransferase, as opposed to the multifunctional CysG enzyme. On the basis of the inferences made above with the similarity of NirE to other SUMT enzymes and the structural topology of CobA, CysG and CbiF being that of a homodimer, one may also predict that NirE will also be homodimeric.



**Figure 3.3: Multiple sequence alignment of known characterised SUMT enzymes**

NirE *Ps. aeruginosa* (279), NirE *Ps. stutzeri* (278), SUMT *Pa. denitrificans* (287) were aligned with CobA *Ps. denitrificans* (280), SirA *B. megatarium* (259), SirA (YlnD) *B. Subtilis* (257), CorA *M. ivanovii* (231), Met1p *S. cerevisiae* (593), *A. thaliana* (369), ZmSUMT1 maize (423), CysG<sup>A</sup> *S. enterica* (242), CysG<sup>A</sup> *E.coli* (255). Sequence identities are evident by the darker coloured boxes, the conserved SAM binding site is shown by the green box (GXGXG). The first portion of the alignment is shown here and the second portion on the next page.



**Figure 3.3 continued.** The second portion of the alignment, with the latter half of the protein sequences.

		280	290	300	310	320	330	340
<i>UPM1pA.thaliana/1-369</i>	265	-DPLFVAENAADPD	TLVVYMG	LGLTLP	SLAQK	LMDHG	LPSDTP	PAVAVERGTTPLQRTVFAELKDFATEIQSAGLVSPT
<i>ZmSUMT1maize/1-423</i>	269	-DPLYVAGNAADPD	TLVVYMG	LSTLPS	LAPKLMKH	G	LPPDTP	PAVAVERGTTTPQQR
<i>SirAB.megaterium/1-259</i>	152	-DAIKWDSLAKGVDT	-LA	IYMG	VRNLPY	ICQQLMKH	G	KTSATPIALIHWTTCADQRTVTGTLG
<i>SirAYnDB.Subtilis/1-257</i>	152	-FEEKWKALATGIDT	-LV	IYMG	IKNVQ	IERK	LENG	RDG
<i>CorAM.ivanovii/1-231</i>	149	-CKKQVGVDFKADT	-I	VILMG	IGNL	AENTAE	IMKH	-KDPETPVCVIENGTMEGQRIITGTLEN
<i>NirEPs.aeruginosa/1-279</i>	168	-LDLDWAGLARQKQT	-LV	FYMG	LGNLAE	I	AARLVEHG	LASDTPAALVSQGTQAGQQVTRGALAE
<i>NirEPs.stutzeri/1-278</i>	168	-LHLPWDALAQGGQT	-LV	FYMG	LASLGE	I	SRNLI	SAGLPVDTPAALIGNGTRE
<i>CysGAS.typhimurium/1-243</i>	151	-ELDWENLAEEKQT	-LV	FYMG	LNQAAT	I	QEKLI	AFQMQADMPVALVENGT
<i>CysGAE.coli/1-256</i>	164	-ELDWENLAEEKQT	-LV	FYMG	LNQAAT	I	QKLI	IEHGMPEMPVAIVENGTAVTQ
<i>SUMTPa.denitrificans/1-287</i>	181	-LDLDWASLADPQTT	-LA	IYMG	AANMAE	I	AREL	IRHGM
<i>CobAPs.denitrificans/1-280</i>	165	PDRINWQGIASGSPV	-I	VY	MAMKH	I	GAL	TANLIAGGRSPDEP
<i>Met1pS.cervisiae/1-268</i>	151	-PIIPEFVESRTT	-V	FLMAL	HHRANV	L	ITGL	LLKHGWDGDVFAAIVERG
		370	380	390	400	410	420	
<i>UPM1pA.thaliana/1-369</i>	354	LWPHCTK	ESSCLVETR	.....	.....	.....	.....	.....
<i>ZmSUMT1maize/1-423</i>	358	FWVESSE	HDALKVQSSY	ANEADDNRF	SWFLVTGQYLRNARSQ	IWEHRFDLVYTET	SQEA	WILNGGW
<i>SirAB.megaterium/1-259</i>	240	RLNWF	EKTGQHYSYA	QEAY	.....	.....	.....	.....
<i>SirAYnDB.Subtilis/1-257</i>	240	KLEWF	ESELKK	-QDLSEAL	.....	.....	.....	.....
<i>CorAM.ivanovii/1-231</i>	231	K	.....	.....	.....	.....	.....	.....
<i>NirEPs.aeruginosa/1-279</i>	256	ERAMAHPSYL	GAGSPVSREAVACA	.....	.....	.....	.....	.....
<i>NirEPs.stutzeri/1-278</i>	256	DRQVEHPARL	-HGDPAPKAEALCN	.....	.....	.....	.....	.....
<i>CysGAS.typhimurium/1-243</i>	236	KLNWF	SNH	.....	.....	.....	.....	.....
<i>CysGAE.coli/1-256</i>	249	KLNWF	SNH	.....	.....	.....	.....	.....
<i>SUMTPa.denitrificans/1-287</i>	269	DCALPQEL	YRPEWRLVAHG	.....	.....	.....	.....	.....
<i>CobAPs.denitrificans/1-280</i>	254	ALDW	IGALDGRKLAADPFANRILRNPA	.....	.....	.....	.....	.....
<i>Met1pS.cervisiae/1-268</i>	236	EKDL	INFDES	RKFVIDEGFREFEVDVDSL	FLKLY	.....	.....	.....

**Table 3.1: A summary of SUMT amino acid homology**

The values represented here are derived from alignments using ClustalW (Figure 3.3) and are given as percentage amino acid sequence identity.

%	ZmSUMT maize	UPM1p <i>A.thal</i>	Met1p <i>S.cere</i>	CorA <i>M.ivan</i>	SirA <i>B.subt</i>	SirA <i>B.meg</i>	CysG <sup>A</sup> <i>S.typhi</i>	CysG <sup>A</sup> <i>E.coli</i>	CobA <i>Ps.deni</i>	NirE <i>Ps.stut</i>	NirE <i>Pa.den</i>
<b>NirE <i>Ps.aeru</i></b>	43	42	30	45	40	42	55	52	41	63	46
<b>NirE <i>Pa.den</i></b>	40	40	29	38	36	37	46	43	37	43	
<b>NirE <i>Ps.stutz</i></b>	44	43	31	41	42	41	56	54	39		
<b>CobA <i>Ps.den</i></b>	38	36	30	40	41	41	44	41			
<b>CysG<sup>A</sup> <i>E.coli</i></b>	48	48	35	41	47	48	92				
<b>CysG<sup>A</sup> <i>S.typhi</i></b>	50	49	35	39	45	48					
<b>SirA <i>B.meg</i></b>	45	44	29	46	62						
<b>SirA <i>B.subt</i></b>	44	41	31	43							
<b>CorA <i>M.ivan</i></b>	44	44	28								
<b>Met1p <i>S.cere</i></b>	34	35									
<b>UPM1p <i>A.thal</i></b>	67										

### 3.3. The overproduction of NirE

#### 3.3.1. Cloning and expression of *nirE* in *E.coli*

Primers designed to *Ps. aeruginosa nirE* were constructed to introduce suitable restriction sites, *NdeI* (TGC ATA TGA ACA CTA CCG TGA TTC CG) and *BamHI* (TCG GAT CCG ATC AGG CGC ATG CGA C), at the 5' and 3' termini of the *nirE* gene respectively. Introduction of these sites allows in frame cloning in to the pET14b (amp<sup>R</sup>) expression vector allowing the encoded protein to be produced as a N-terminal poly-Histidine tagged fusion protein, conveniently allowing protein overproduction and affinity purification by metal chelate chromatography

Amplification of *nirE* was performed by PCR using the respective designed primers and HotStar*Taq* DNA polymerase (Qiagen). Conditions for the amplification can be found in Section 2.2.8.1.1. The PCR product of predicted size was separated by agarose gel electrophoresis, extracted and cloned into pGEMT-Easy (pVP01); *nirE* was then subcloned into pET14b expression vector (pVP02) using the *NdeI*, *BamHI* restriction sites, behind a poly-histidine encoding sequence, under the control of a T7 promoter. Refer to Chapter 2, Section 2.2 for protocols.

To facilitate recombinant protein overproduction, the pET14b-*nirE* plasmid was transformed into *E. coli* BL21 (DE3) pLysS strains, in which the T7 polymerase is induced by the addition of IPTG, to a final concentration of 0.4 mM in cultures grown in LB medium at 37 °C (plus antibiotics) with aeration. Cultures were further incubated at 37 °C or 16 °C with aeration, before harvesting by centrifugation (10 000 x g, 4 °C). The bacterial pellets were primarily resuspended in binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 8.0) before protein extraction and purification (Section 2.3.3).

#### 3.3.2. Cloning and expression of *nirE* in *Ps. aeruginosa*

Since the pET14b T7 promoter is not recognised by *Ps. aeruginosa* pUCP19-*NcoI*, under the control of a *lac* promoter, was used as compatible plasmid for expression in *Ps. aeruginosa* (Chuanchuen, Narasaki et al. 2002). *nirE* was subcloned into pUCP19-*NcoI* (pVP17), from pET14b using the restriction sites *NcoI* and *BamHI*, sub-cloning the gene with the Histag to allow recombinant protein purification.

To facilitate recombinant protein overproduction, the pUCP19-Nco-*nirE* (pVP17) plasmid was transformed into wild type *Ps. aeruginosa* (PAO1), in which the *lac* promoter is induced by the addition of IPTG. A final concentration of 1 mM IPTG was added in cultures grown in LB or SLB medium at 37 °C (plus carb 500 mg/l) with aeration. Cultures were further incubated at 16 °C with aeration, before harvesting by centrifugation (10 000 x g, 4 °C). Anaerobic cultures were grown in the same way but on a nitrate-containing medium (25 mM) with mineral oil to deplete oxygen and without aeration. The bacterial pellets were resuspended in binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 8.0) before protein extraction and purification as detailed in Section 2.3.3.

### **3.3.3. Purification of recombinant NirE from *E. coli* and *Ps. aeruginosa*.**

Overproduced recombinant NirE from *E. coli* and *Ps. aeruginosa* was purified as described in the pET system manual (Novagen). Typically, prior to purification, the crude cell fractions were sonicated and centrifuged (10 000 x g, 4°C). The resultant supernatants containing soluble protein were then loaded onto separate charged His-Bind columns. The columns were washed with binding buffer, and wash buffers of increasing imidazole concentration (50 mM then 100 mM), before eluting the protein at 400 mM imidazole. The presence of protein in the fractions was detected by the BioRad protein assay and protein integrity was verified by SDS-PAGE.

### **3.3.4. SDS Analysis of recombinant NirE from *E. coli* and *Ps. aeruginosa***

SDS-PAGE gels containing 12 % polyacrylamide were used to analyse purified NirE, providing sufficient resolution for NirE that is predicted to be 31.9 kDa.

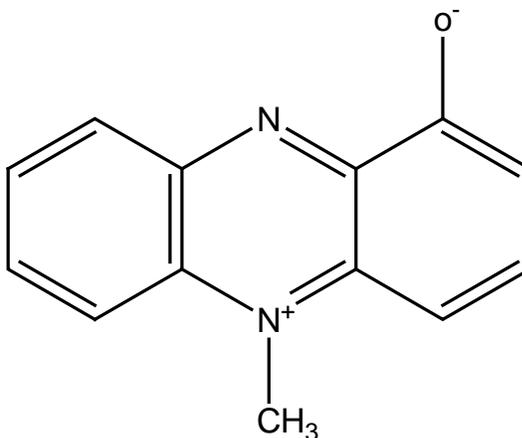
NirE was purified to homogeneity and is soluble in both *E. coli* (44 mg/L) and *Ps. aeruginosa* (34.5 mg/L) in 50 mM Tris pH8, 0.5M NaCl. Protein integrity was verified by SDS-PAGE of the final eluted fractions from both organisms. The gene for *nirE* is 839 bases therefore the protein is predicted to be 29.7 kDa. With the His-tag we see this at 31.9 kDa on the SDS-PAGE gels in Figure 3.5 and 3.6. The yield of protein obtained by incubation at 37 °C (32 mg/L) after the addition of IPTG was less than that obtained at 16 °C (44 mg/L). Subsequent experiments were performed by incubation at the lower temperature. The yield was lower in *Ps. aeruginosa* and to enhance overproduction, *nirE* was overexpressed in *Ps. aeruginosa* grown on Super LB broth, and was compared that grown on LB, to provide more nutrient

media as induction for long periods is not possible in this organism due to the production of pyocyanin (Whooley 1982). Pyocyanin is a toxic product produced when growth factors are limiting for the organism and thus is a limiting factor in the overproduction of proteins in *Ps. aeruginosa*.

Each culture was grown for the same time, induced and harvested and purified together in exactly the same manner for comparison. However, this made no significant difference to NirE yields with both producing approximately 30-34 mg/L of NirE as observed in Figure 3.7. Further analysis of NirE by gel filtration chromatography shows that NirE exists as a homodimer in its native form, as predicted. Interestingly, the concentrated NirE fractions exhibited red/ burgundy hues. The UV-Vis spectrum of this is shown in Figure 3.8 and is consistent with the presence of trimethylpyrrocorphin, a product of the overmethylation of UIII observed in the soluble cell extract of other SUMTs, such as CysG (Warren, Stolowich et al. 1990). It appears that NirE uniquely retains this even through purification, being quite tightly bound whereas other SUMTs do not exhibit this phenomenon. Other properties exhibited by this pigment are discussed in section 3.35.

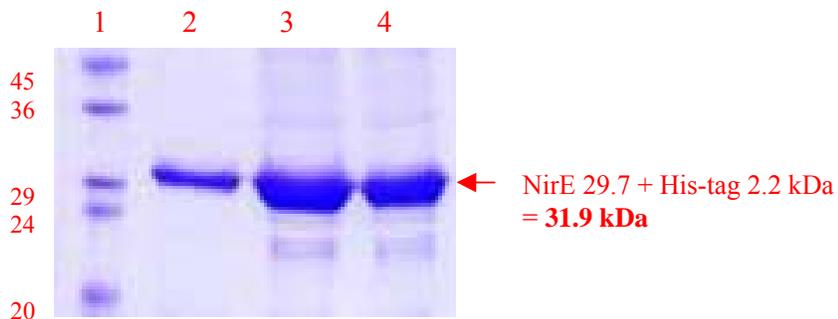
**Figure 3.4: The structure of pyocyanin**

Pyocyanin is a phenazine pigment produced by *Ps. aeruginosa*, when growth factors are limiting, and is responsible for the production of reactive oxygen species (Whooley 1982).

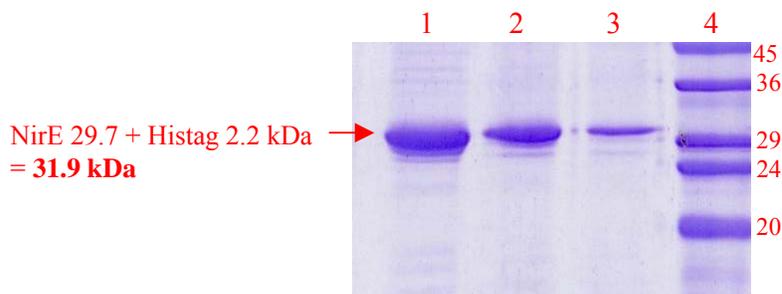


**Figure 3.5: SDS-PAGE of NirE from *E.coli*.**

Lane 1 contains the molecular mass marker as indicated (kDa). Lanes 2, 3, and 4 represent each 2 ml NirE protein fraction eluted from the His-bind column, which is clearly evident by bands at 31.9 kDa (NirE is 29.7 kDa plus N-terminal poly histidines at 2.2 kDa). These results are from 1L cultures grown at 37 °C and induced for protein expression at 16 °C, overnight. NirE = 44 mg/L.

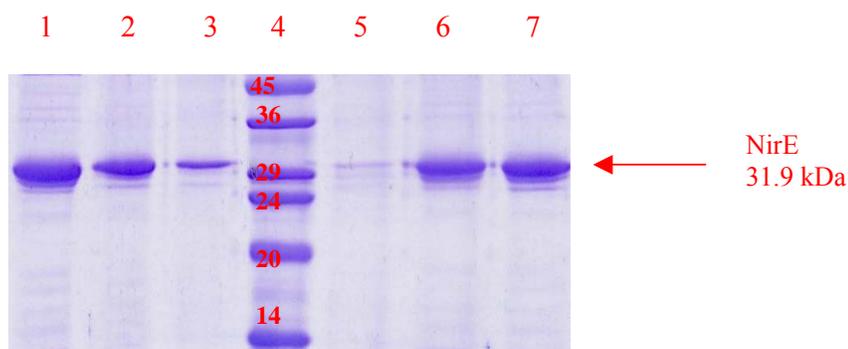
**Figure 3.6: SDS-PAGE of NirE from *Ps. aeruginosa*.**

Lanes 1, 2 and 3 represent each 2 ml NirE protein fraction eluted from the His-bind column, which is clearly evident by bands at 31.9 kDa (NirE is 29.7 kDa plus N-terminal poly histidines at 2.2 kDa). Lane 4 contains the molecular mass marker as indicated (kDa). These results are from 1L cultures grown at 37 °C and induced for protein expression at 16 °C, overnight. NirE = 34.5 mg/L.

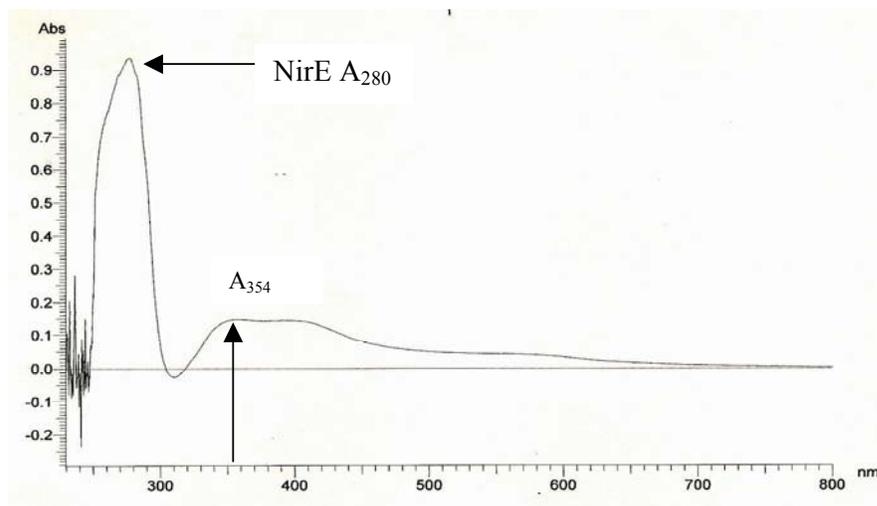


**Figure 3.7: SDS-PAGE of NirE from *Ps. aeruginosa***

Lanes 1, 2 and 3 represent each 2 ml NirE protein fraction eluted from the His-bind column of NirE overproduced in LB. Lane 4 contains the molecular mass marker as indicated (kDa). Lanes 5, 6, 7 represent each 2 ml NirE protein fraction eluted from the His-bind column of NirE overproduced in SLB. These results are from 1L cultures grown at 37 °C and induced for protein expression at 16 °C, overnight. We can see that there is no significant difference in the amount of protein produced in LB or SLB. NirE 30-34 mg/L for both LB and SLB

**Figure 3.8: UV-VIS Spectra of NirE with porphyrinoid material**

Concentrated NirE fractions have red / burgundy hues. The UV-Vis spectrum is consistent with the presence of trimethylpyrrocorphin with absorbance maxima at 354 nm. This is a product of the overmethylation of UIII, normally observed in the soluble cell extract of other SUMTs, but here it is tightly bound to purified NirE.



### 3.3.5. The accumulation of a red pigment, trimethylpyrrocorphin

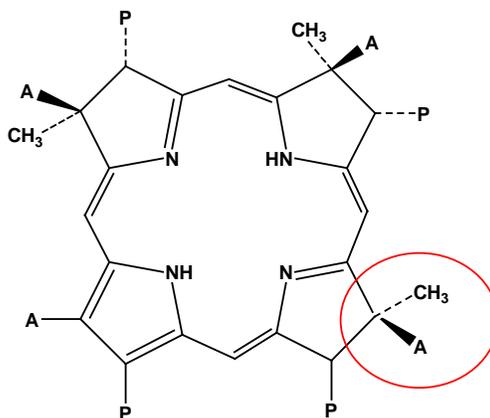
A property common to SUMT activity is the accumulation of a red pigment due to the presence of a mixture of porphyrinoid material mainly trimethylpyrrocorphin (Figure 3.9). This is observed when CobA, CysG, Met1p, ZmSUMT1 and UPM1p are overexpressed as recombinant proteins in *E. coli* and has absorbance maxima at 354nm (Warren, Stolowich et al. 1990; Raux, McVeigh et al. 1999; Fan, Wang et al. 2006). NMR analysis shows it to be a result of overmethylation of PC-2 by addition of a third SAM-derived methyl group at position C-12. When illuminated by UV light, cells and the isolated cell fractions also fluoresce a characteristic bright pink / red colour. There is no physiological role for it and no implication for this material in the biosynthesis of tetrapyrroles. It is likely to be an artifact of non-physiological conditions that would otherwise not occur *in vivo* (Warren, Bolt et al. 1994). This property means that SUMT enzyme activity can be easily detected. Additionally, *cobA* or *cysG* genes can be employed in a wide range of cell types as a fluorescence-based method for the selection of recombinant plasmids as the replacement or disruption of these genes with other fragments of DNA results in loss of enzymatic activity and nonfluorescent cells. Thus inserts are detected in clones lacking fluorescence (Roessner and Scott 1995; Roessner 2002). Interestingly, the red pigment was not observed with CorA suggesting greater flux through the pathway (Robin, Blanche et al. 1991; Fan, Wang et al. 2006).

In the case of NirE it was noted that cells cloned with *nirE* exhibit fluorescence under UV light, as shown in Figure 3.10. During purification, isolated cell fractions, especially the supernatant containing soluble NirE prior to purification was a deep red/brown colour and exhibited bright red / pink coloured fluorescent activity under a UV trans-illuminator. This has been shown in Figure 3.11, where the dark colour of the supernatant is clearly visible on both the column and the eppendorf with red/pink fluorescence when illuminated by UV light. This is consistent with the SUMT activity predicted for NirE, indicating the presence of porphyrinoid pigments. The group of Layer *et al* recently analysed this pigment and confirmed that it is indeed trimethylpyrrocorphin (FEBS Journal, 2009, in press). Further, the presence of this pigment still bound to concentrated NirE may suggest a greater flux through the haem *d*<sub>1</sub> pathway, and a lack of inhibition of NirE by substrate (UIII) or product (SAH). These characteristics being attributed to most SUMT isolates

further suggest NirE is an active UIII methyltransferase and performs methylation of UIII to PC-2 *in vivo*.

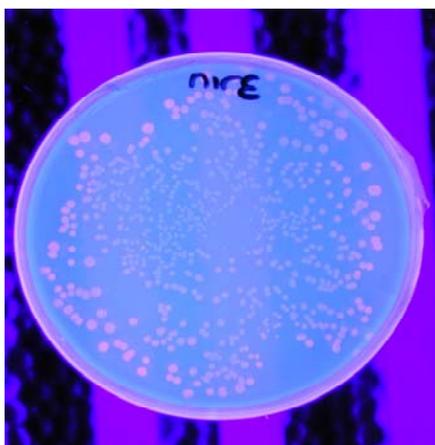
**Figure 3.9: The structure of trimethylpyrrocorphin**

SUMT enzymes sequentially methylate UIII at C2 and C7 yielding PC-2. Further incubation of SUMT enzymes with PC-2 and SAM result in an additional methylation at C12 to produce 2,7,12 - trimethylpyrrocorphin, a non-physiological product.



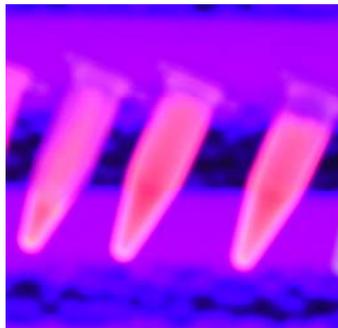
**Figure 3.10: NirE cells containing trimethylpyrrocorphin fluoresce red under UV light.**

NirE pET14b transformed into BL21 (DE3) pLysS, plated with amp and IPTG and incubated at 37 °C. When illuminated by UV light the fluorescent red colour can be clearly seen in clones with *nirE*, a phenomena characteristic of *cobA* and most SUMT enzymes (Warren, Stolowich et al. 1990), indicating NirE is a SUMT.



**Figure 3.11: The characteristic red pigment of SUMTs shown by NirE.**

In **A**, the metal chelate column has been changed from blue to brown by the addition of the supernatant from cells overproducing NirE. In **B**, the red pigment is clearly visible in the supernatant and in **C**, the UV fluorescence of trimethylpyrrocorphin is shown. These entire characteristics are attributed to SUMTs and signify that NirE is a SUMT.

**A****B****C**

### 3.4. Complementation of *Salmonella enterica cysG metH* mutant strains by *Ps. aeruginosa* NirE

To investigate the role played by NirE in the tetrapyrrole biosynthetic pathway, a plasmid containing *nirE* (pVP16) was used to complement a *S. enterica cysG metH* mutant (Table 2.1, provided by Dr Deery). *S. enterica cysG metE* (AR3612) is a CysG cysteine auxotrophic strain, unable to produce sirohaem or synthesise cysteine as shown in Chapter 1. It is deficient only in its methyltransferase activity having no CysG<sup>A</sup> activity but retaining CysG<sup>B</sup> activity as the host contains a gene having the same phenotype as *cysG<sup>B</sup>* (*cbiK*) (Raux, Lanois et al. 1996). *S. enterica metE* cells are reliant upon their alternative B<sub>12</sub>-dependent methionine synthase, encoded by *metH*, for the biosynthesis of methionine. Therefore it lacks the ability to catalyse the transformation of UIII to PC-2.

On the basis of complementation with *cysG* mutants from *E. coli* (*cysG* 302Δa) and *S. enterica* (*cysG metE* AR3612), MET1p was shown to confer SUMT activity, as it was able to complement the cysteine auxotrophy of the *S. enterica* mutant, but not the *E. coli* strain. It has been shown that *S. enterica cysG* mutants are blocked in cobalamin and sirohaem synthesis thus implicating its role in both, as opposed to *P. denitrificans cobA* mutants that retain less than 2 % of their parent activity. They continue to grow on minimal media without cysteine, suggesting *Ps. denitrificans* does make sirohaem via another SUMT, CysG, and the presence of more than one SUMT. This *S. enterica* mutant is deficient in only its CysG<sup>A</sup> whereas the *E. coli* strain is deficient in all CysG. Furthermore, using a different strain, it was shown that ZmSUMT1 and UPM1p were able to complement *E. coli* CysG<sup>A</sup> mutants by overcoming cysteine autotrophy on minimal media. (Crouzet, Cauchois et al. 1990)

Complementation studies were conducted by transforming into *S. enterica cysG metE* (AR3612) competent cells. The resultant colonies were then restreaked on to minimal media plates made with 50 µg/ml of both methionine and cysteine; with cysteine and without methionine; without cysteine and with methionine; without cysteine and methionine; plus antibiotics as appropriate. The plates were then incubated at 37 °C for 24-48 hrs, until bacterial colonies were evident in the plates. The results have been summarised in Table 3.2.

*nirEpUCPNde* (pVP16) was used to confirm its ability to complement SUMT activity; *nirFpUCPNde* (pVP18) was used to check for any methyltransferase activity as it has been reported that NirF has 21 % sequence identity to NirE in the N-terminal 100 amino acids, but the aligned SAM binding site of NirE (GAGPG) is different in NirF (LRGSG) (Kawasaki, Arai et al. 1997); *nirFDLGHJpUCPNde* (pVP63) was used to confirm again that NirE is the only SUMT of this operon; pUCPNde and *cobA* respectively acted as negative and positive control.

The results obtained show for the first time that *nirE* is able to complement the cysteine auxotrophy of the *S. enterica* mutant (*cysG metE* AR3612) by its capacity to grow in the absence of exogenous cysteine. Thus it is able to complement the lack of CysG<sup>A</sup> in the sirohaem and vitamin B<sub>12</sub> pathways and confers UIII methyltransferase activity in the transformation of UIII to PC-2. Additionally, the requirement of NirE for methionine to perform this reaction is also evident by the lack of growth in the absence of both methionine and cysteine as it is unable to overcome this. As predicted *cobA* gave similar results to *nirE*; pUCPNde, *nirFDLGHJ* and *nirF* were unable to complement the cysteine auxotrophy of the *S. enterica* mutant. This confirms that *nirF* confers no methyltransferase activity and that *nirE* is the only SUMT as part of the *nir* operon.

**Table 3.2: Summary of results obtained from complementation of *S. enterica* *cysG metE* (AR3612)**

Complementation was performed using *nirE*, *cobA*, *nirF* and *nirFDLGHJ*. +++ indicates good growth and normal colony size after 24 hrs; ++ indicates slow cell growth and small colonies after 24 hrs but normal cell size after 48 hrs; + indicates very poor cell growth and small colony size after 48 hrs; - indicates a negative response, no colonies

Plate conditions	<i>cobA</i> pET14b	<i>nirE</i> pUCPNde	pUCPNde	<i>nirF</i> pUCPNde	<i>NirFDLGHJ</i> pUCPNde
+ Cysteine + Methionine	+++	+++	+++	+++	+++
+ Cysteine - Methionine	++	++	++	++	++
- Cysteine + Methionine	++	++	-	-	-
- Cysteine - Methionine	+	+	-	-	-

### 3.5. Biochemical complementation of NirE

NirE, thus far, was able to complement the *S. enterica cysG methH* mutant. It was able to restore sirohaem and vitamin B<sub>12</sub> biosynthesis in this strain, but there remained ambiguity as to whether it would be possible to synthesise experimentally PC-2 *in vitro* with NirE and thereafter SHC using SirC. It is predicted that NirE is able to synthesise PC-2 with the same stereo-regio specificity as the sirohaem, or vitamin B<sub>12</sub> pathways, that is then converted to SHC. The production of both PC-2 and SHC under anaerobic conditions is fully characterised and been performed using CobA (Leech, Raux-Deery et al. 2002; Raux, Leech et al. 2003). The first point of interest was to see if NirE could make both PC-2 and SHC *in situ* from soluble crude cell extracts of the enzymes required to make PC-2 and SHC, in the same way that CobA is able to do so. On the basis of this, a SHC coupled assay can be used to look at the activity of NirE in the presence or absence of defined co-factors and substrates.

The production of PC-2 and SHC can be monitored by marked differences in the spectra of the two compounds and their distinctive colouring; PC-2 is yellow and exhibits a broad spectrum, with no distinct absorption maxima; SHC is purple and has UV-Vis absorption maxima at 376 and 590 nm. The differences arise due to the increased level of conjugation in SHC, around the macrocyclic ring by the addition of a double bond. This appearance of SHC in the transformation of PC-2 to SHC can therefore be easily followed by spectroscopy as shown in Figure 3.12 by the increase in absorbance at 376 nm.

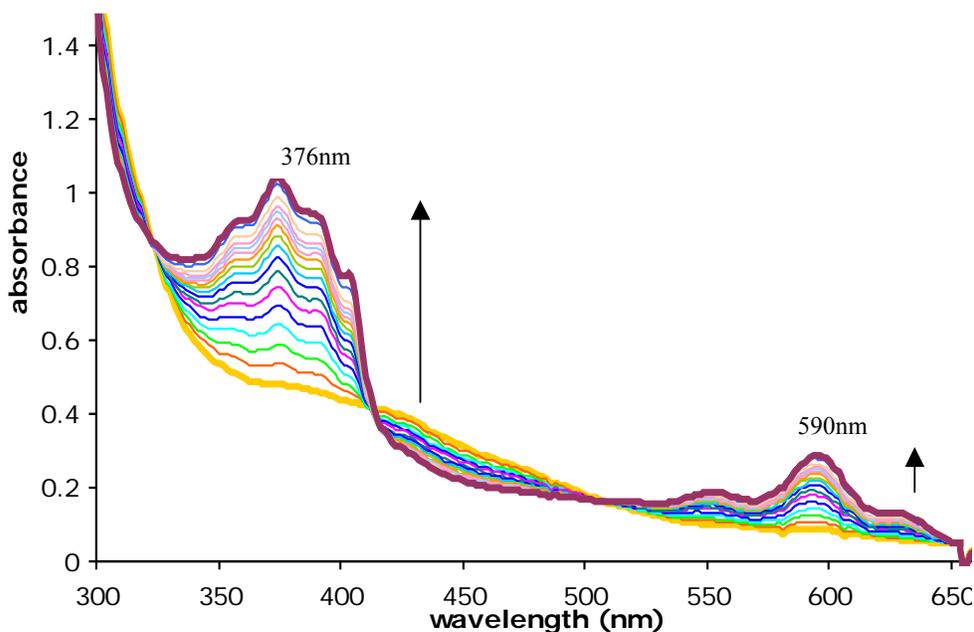
#### 3.5.1. Production of PC-2 and SHC *in situ* using NirE

PC-2 and SHC were produced from ALA, using CobA and NirE under a nitrogen atmosphere, in a glove box with less than 2 ppm oxygen as described in Section 2.4.2. The resultant UV-Vis spectra for PC- 2 and SHC produced with CobA are shown in Figure 3.13. PC-2 and SHC produced with NirE is shown in Figures 3.14 and 3.15 respectively. Briefly, crude soluble cell extracts containing overproduced HemB, HemC, HemD and either CobA or NirE, were added to 10.0 mg ALA and 10 mg of SAM, in a total volume of 40 ml of 50 mM Tris-HCl buffer, pH 8.0. Sirohydrochlorin was generated using the same enzyme cocktail as above for PC-2 but included SirC and 13 mg of NAD<sup>+</sup>. The reaction mixture was left overnight at 22 °C to allow it to reach completion, and the resulting tetrapyrroles filtered before

spectroscopic analysis. It was notable that whilst CobA was seen to turn over a small amount of the substrate after an hour, NirE had not, and took longer to begin turning over the substrate. Also, the spectra for NirE PC-2 and SHC can be seen to have a great deal of precipitation, even after filtration presumably due to the presence of NirE and its relative instability. Furthermore, when left for a further 24 hrs, PC-2 was turning slightly orange and an UV-Vis scan of the filtered reaction mixture showed that PC-2 was slowly being oxidized due to its high sensitivity to oxygen. SHC was seen to be turning slightly blue after 24 hrs; again an UV-Vis scan showed that it too had a less well-defined spectrum. Therefore, we report here for the first time the formation of PC-2 and SHC with the use of NirE as a SUMT in the reaction.

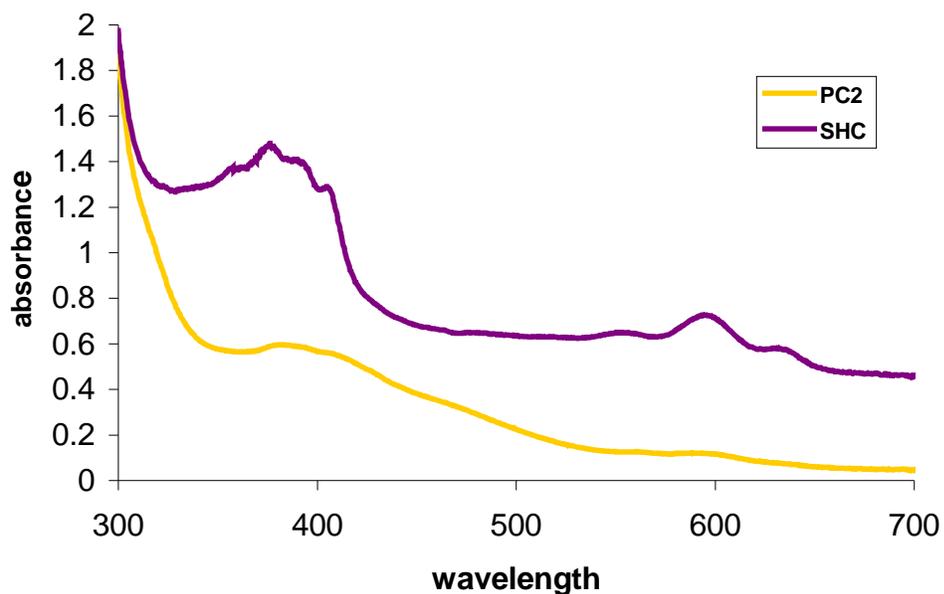
**Figure 3.12: Spectra showing the transformation of PC-2 to SHC by CobA.**

The conversion of PC-2 to SHC was followed by UV-Vis spectroscopy over time. PC-2 is shown as the thick yellow band, mostly at the bottom and SHC is shown by the thick purple spectra, mainly at the top. The generation of SHC is seen by the increase in the absorption maxima at 376 nm and 590 nm shown by the arrows, characteristic of SHC



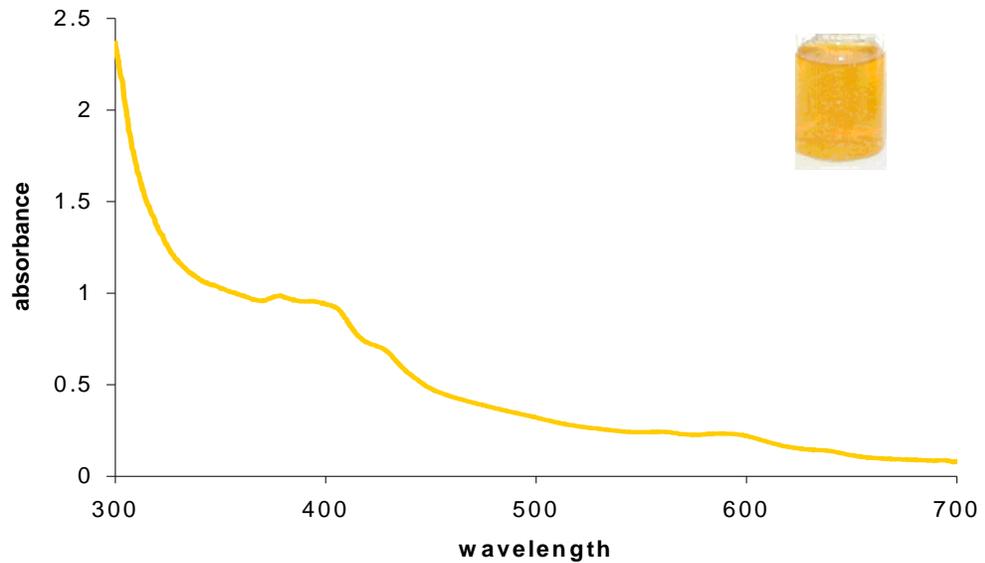
**Figure 3.13: *In situ* production of PC-2 and SHC with CobA.**

UV-Vis spectra from incubation of CobA with soluble crude cell extracts of the enzymes required for the synthesis of PC-2 and SHC from ALA. Typical spectra obtained for PC-2 is seen in yellow with a band 350 nm to 430 nm. Spectra characteristic of SHC is shown in purple, with absorption maxima at 376nm and 590 nm (also at 388 nm and 400 nm). As expected, the mixtures were yellow and purple for PC-2 and SHC respectively.



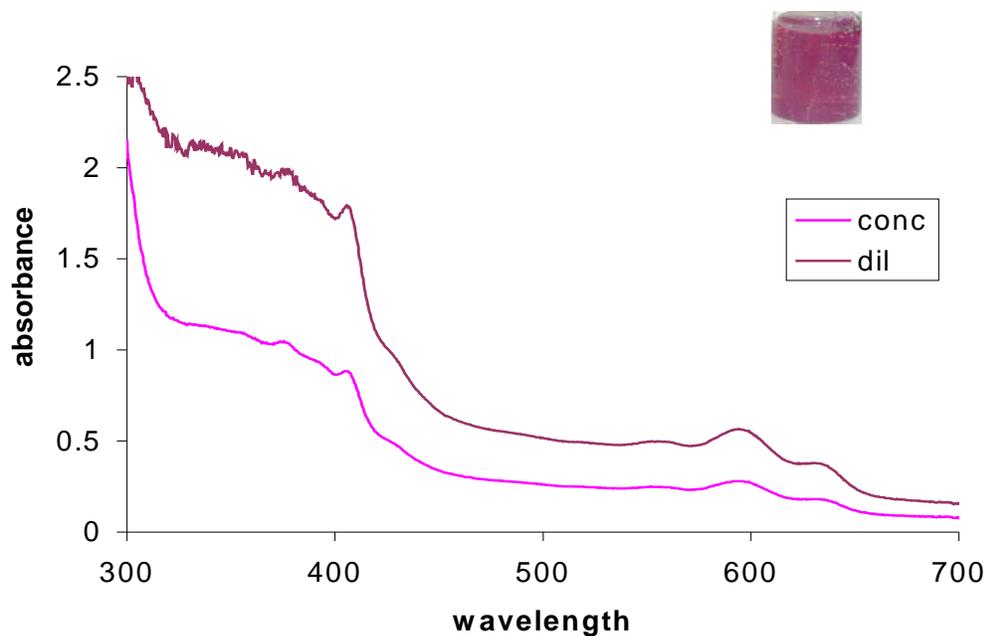
**Figure 3.14: *In situ* production of PC-2 with NirE.**

UV-Vis spectra obtained from incubation of NirE with the soluble crude cell extracts of the enzymes required for the synthesis of PC-2. The spectra for PC-2 can be seen with a band at 350 to 430nm. The solution was observed to be yellow, as expected, indicating that NirE is able to make PC-2. In inset picture shows the colour of PC-2 after filtration.



**Figure 3.15: *In situ* production of SHC with NirE.**

The UV-Vis spectra below are obtained from incubation of NirE along with the soluble crude cell extracts of the enzymes required for the synthesis of SHC. The spectrum for SHC is shown in purple with absorption maxima at 376 nm and 590 nm. The solution was observed to be purple, as expected, indicating that NirE is also able to make SHC. The inset picture shows the colour of SHC after filtration.



### 3.6. SHC coupled NirE assay

It has now been shown that NirE is able to make PC-2 and SHC *in situ*. This can be utilised to look at the activity of NirE in the presence or absence of SAM and NAD<sup>+</sup>, and with or without NirE under strictly anaerobic conditions. This will give us an indication of the dependency of the transformation on NirE our SUMT and further the dependency of NirE on co-factors, mainly SAM.

SHC was used as an indicator of SUMT activity instead of PC-2. The reasons for coupling SUMT activity to SHC are many. SHC is less sensitive to oxygen than PC-2 that is very sensitive to an oxygen concentration of greater than 2 ppm, monitoring PC-2 will be more difficult as it may spontaneously oxidise. SHC is more easily monitored by UV-Vis spectroscopy as the spectra for SHC is more defined than that of PC-2, as seen in Figure 3.12.

The assay was performed using purified NirE, and HemB, HemC, HemD, and SirC. NirE was overproduced and purified separately, as before. A plasmid containing all four genes required to produce SHC from ALA, *hemB*, *hemC*, *hemD*, and *sirC*, pET14b-BCDC, was similarly overexpressed and the resultant overproduced recombinant proteins were purified in the same fashion. All proteins and cofactors were added in excess, except for NirE, which was added last. The reaction was followed at timed intervals and the change in absorbance at 376 nm was followed to monitor the generation of SHC since SHC has an absorption maximum at 376 nm. The concentration of NirE was varied at 9 µg/ml and 4.5 µg/ml total in the reaction. The effect on the absence of NirE, SAM and of NAD<sup>+</sup> was also monitored similarly. The specific activity was calculated using the extinction coefficient of SHC as  $2.4 \times 10^5 \text{ M}^{-1}\text{cm}^{-1}$ . For a protocol see Section 2.4.5.

From the data obtained by the SHC coupled assay it is clear that NirE is performing the transformation of UIII to PC-2. The presence of SAM, NAD<sup>+</sup> and HemB, HemC, HemD and SirC along with a defined concentrations of NirE catalyses the formation of SHC from ALA, as shown in Figure 3.16. The assay was performed with a NirE concentration of 2.25 µg/ml, 4.5 µg/ml and 9 µg/ml. The calculated turnover rate at 2.25 µg/ml NirE is 24/h (data not shown); at 4.5 µg/ml NirE is 27/h and at 9 µg/ml NirE is 40/h showing a concentration dependent effect on the speed of the reaction as expected, seen in Figure 3.17. The absence of NirE halts the

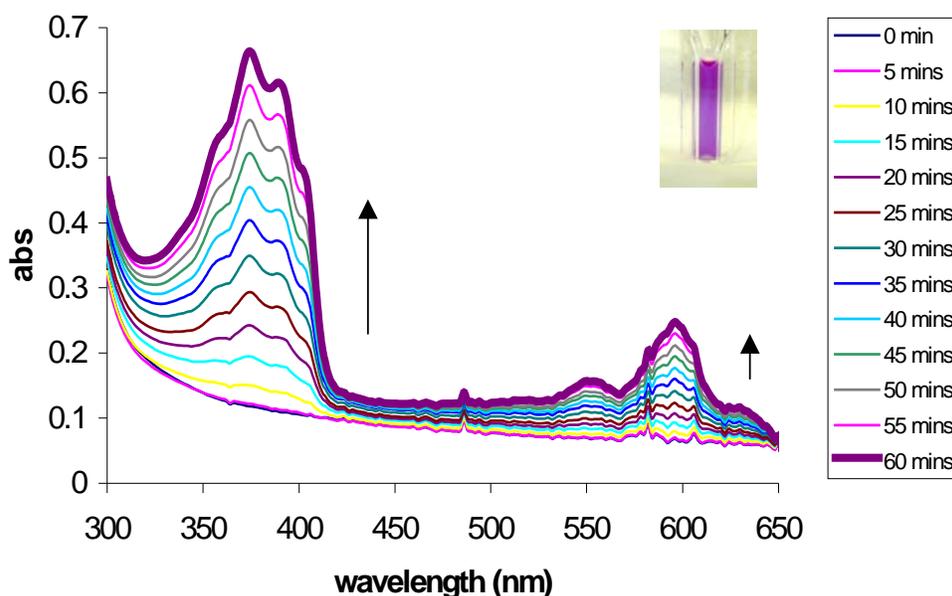
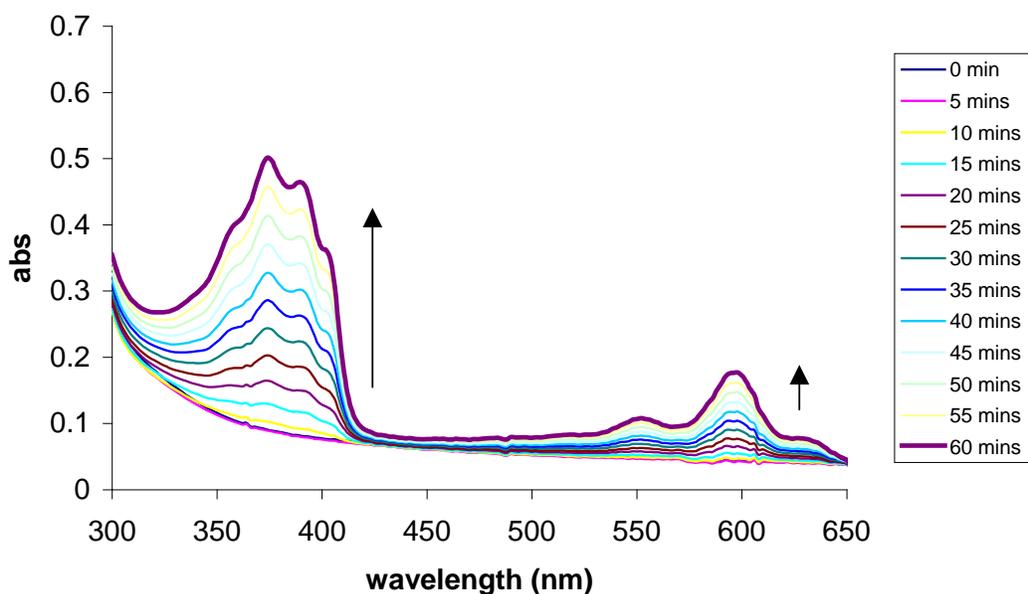
reaction indicating a reliance on NirE for SUMT activity in the assay, as no colour change is observed. All subsequent reactions were performed with 4.5 µg/ml NirE.

The absence of SAM from the reaction demonstrates NirE is dependent on SAM to perform this reaction as with other SUMT enzymes. Omitting SAM from the reaction mixture does not form SHC because the substrate for this reaction PC-2 is not being formed by the methylation of UIII by NirE, this is shown in Figure 3.18.

The absence of NAD<sup>+</sup> from the reaction mixture also inhibits SHC production, but in this case we clearly have the formation of PC-2 by NirE. The reaction cannot proceed further from PC-2 to SHC as SirC activity is perturbed by the lack of NAD<sup>+</sup>. The formation of PC-2 is shown in Figure 3.19, and the yellow colour of PC-2 is seen also.

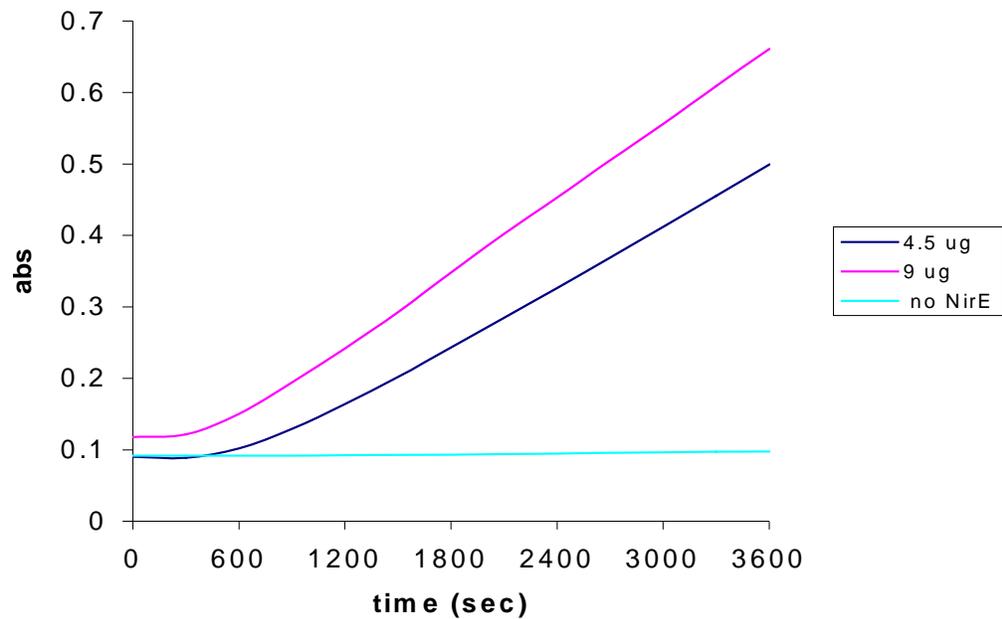
**Figure 3.16: UV-Vis Spectra showing the *in vitro* generation of SHC by NirE**

NirE is able to make SHC *in vitro*, under anaerobic conditions. The reaction conditions remained as stated with the exception of the NirE concentration. In **(A)** NirE is added to a final concentration of 9  $\mu\text{g/ml}$  and in **(B)** NirE is added to a final concentration of 4.5  $\mu\text{g/ml}$ . Each spectrum denotes a wavelength scan at 5 min intervals. The arrows show the direction of change in the spectra. The picture inset shows SHC formed by NirE, as the solution changes from clear to purple.

**A. 9.0  $\mu\text{g}$  NirE****B. 4.5  $\mu\text{g}$  NirE**

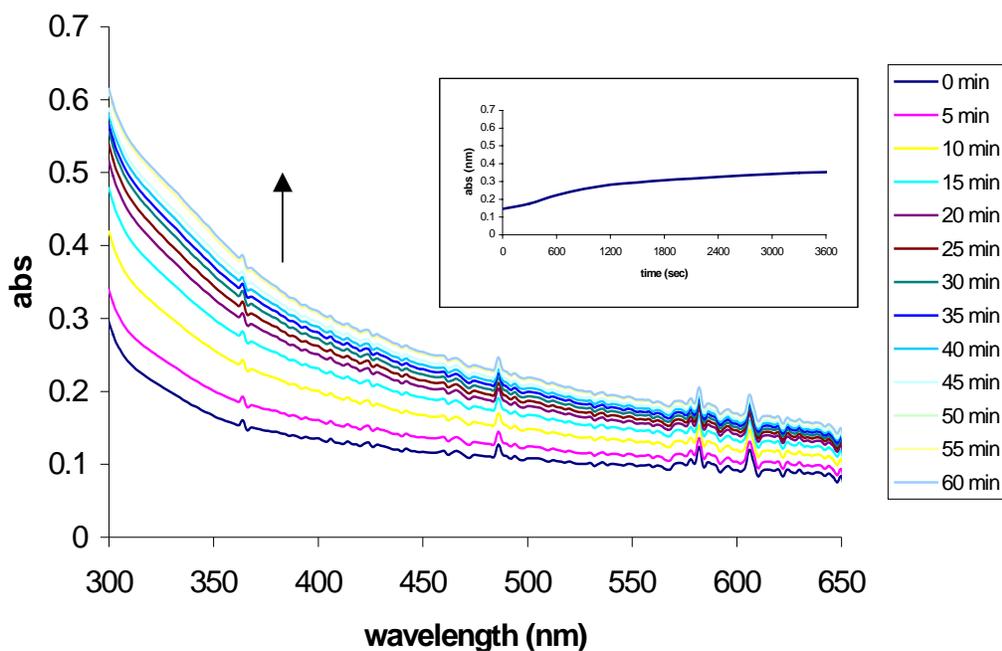
**Figure 3.17: The rate of generation of SHC by NirE**

The rate of formation of SHC by NirE is shown at NirE concentrations of 9  $\mu\text{g/ml}$  and 4.5  $\mu\text{g/ml}$  as well the absence of NirE. The absorbance at 376 nm was followed over time plotted below. The specific activity was calculated accordingly. The reaction is dependent on NirE to transform UIII to PC-2 as the absence NirE results in the inhibition of SHC formation.



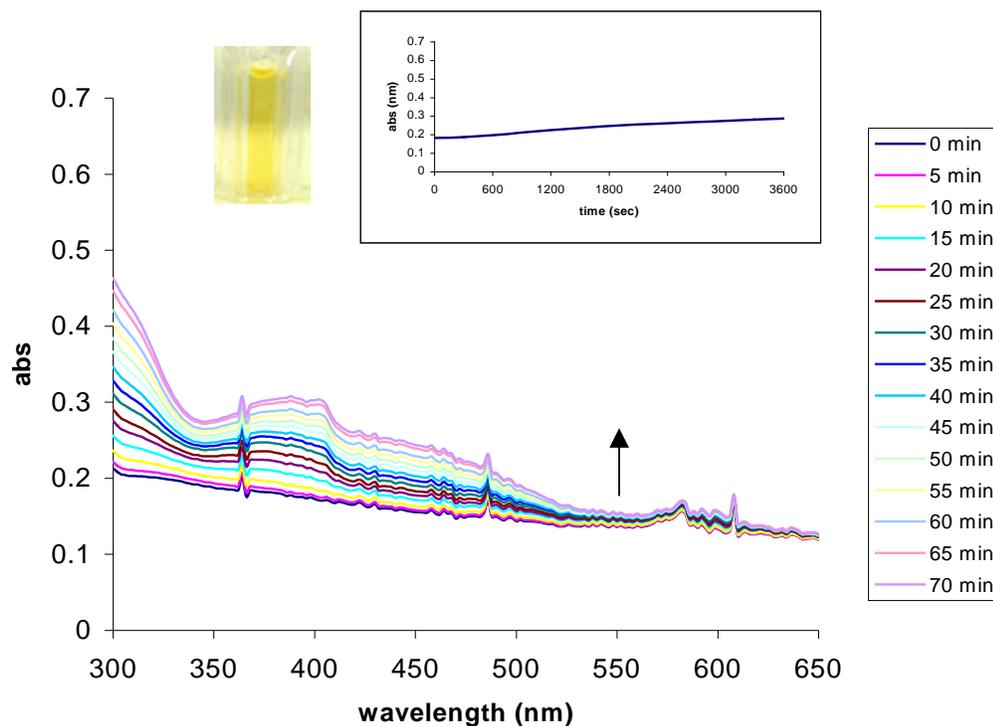
**Figure 3.18: UV-Vis Spectra of reaction mixture without SAM**

The reaction conditions remained as stated with the concentration of NirE at 4.5  $\mu\text{g/ml}$  with the exception of there being no SAM added. SHC is not formed and the resultant UV-Vis spectra, show no NirE activity (inset) indicating that NirE is a SAM dependent UIII methyltransferase. The small observed increase is likely due to precipitation as the incubation turned slightly cloudy. The solution remained colourless, as is UIII.



**Figure 3.19: UV-Vis Spectra of reaction mixture without NAD<sup>+</sup>**

The reaction conditions remain as stated with the concentration of NirE at 4.5  $\mu\text{g/ml}$  with the exception of there being no NAD<sup>+</sup> added. As SirC is NAD<sup>+</sup> dependent SHC is not formed and the resultant UV-Vis spectra, shows no NirE activity (inset). After a period of time the cuvette turned yellow indicating the presence of PC-2 from the picture and from its spectra. This confirms that NirE is able to synthesise PC-2 but since there is no NAD, SirC is unable to take the reaction to completion.



### 3.7. Conclusion

NirE is well documented to be an *S*-adenosyl-L-methionine Uroporphyrinogen III methyltransferase in literature, but it has never been shown experimentally to perform the transformation of UIII to PC-2. Initially, *nirE* was cloned to allow expression in both *E. coli* and *Ps. aeruginosa*. Then it was investigated for complementation of a *S. enterica cysG methH* knockout strain. The next point of interest was to biochemically complement CobA in the pathway to SHC via PC-2 under anaerobic conditions. It was investigated if NirE can perform the same reaction using crude cell extracts as has been performed on many occasions by CobA. The dependency of NirE on cofactors was also investigated with a SHC coupled SUMT assay.

This thesis shows for the first time that NirE exhibits the same characteristics during purification that are attributed to other SUMT enzymes such as CobA and CysG<sup>A</sup> (Warren, Stolowich et al. 1990). The supernatant was deep red in colour and exhibited fluorescent activity under a UV trans-illuminator indicating the presence of tetrapyrrole pigments, including trimethylpyrrocorphin (Figure 3.9). NirE is unique however in that it remains tightly bound to the porphyrinoid even through purification and may suggest a greater flux through the haem *d*<sub>1</sub> pathway, and a lack of inhibition of NirE by substrate (UIII) or product (SAH). From the observed characteristics of NirE in the purification process, we can implicate NirE as our Uroporphyrinogen III methyltransferase, catalysing the first step, the transformation of UIII to PC-2.

Since *Ps. aeruginosa* does not recognise the T7 promoter, a new system was formulated to facilitate recombinant protein overproduction in *Ps. aeruginosa*. It was expected that NirE would be well overproduced in this organism as it is being homologously expressed, and the vector contains a weaker promoter. It was expressed, but not as well as in *E. coli* and likely to be due to the tight regulation of the expression of these genes in *Ps. aeruginosa*. Therefore, despite sharing similarities with other SUMT enzymes, *nirE* expression is tightly regulated and evolved to respond to different regulatory signals than those of the sirohaem and haem *b* pathways. The *nir* gene operons in *Ps. aeruginosa* and *Ps. stutzeri* have been shown to have recognition motifs for the anaerobic regulator ANR in their promoter regions (Arai, Igarashi et al. 1994). In *Pa. denitrificans* the *nir* cluster is

under the control of a NNR transcriptional activator which responds to NO (Van Spanning, De Boer et al. 1995; Saunders, Houben et al. 1999). Conversely, the sirohaem pathway is regulated by nitrogen and is independent of oxygen. Thus, the expression of *nirE* is clearly a well-regulated process in *Ps. aeruginosa*.

NirE is also shown for the first time to complement CysG<sup>A</sup> activity in a *S. enterica* *cysG metH* mutant. This strain transformed with *nirE* from *Ps. aeruginosa* was able to restore sirohaem and methionine biosynthesis and thus complement the cysteine and methionine auxotrophy. This implies the role of NirE is of SUMT activity and the transformation of UIII to PC-2 and is able to complement the sirohaem and cobalamin pathways. Additionally, *nirF* was also postulated to be a methyltransferase having 21% 100 amino acid identity to NirE at the N-terminal (Kawasaki, Arai et al. 1997), this did not complement the mutant strain as expected, since there is no SAM binding site with NirF. The same result was observed with a *nirFDLGHJ* plasmid transformed into the complementation strain, indicating that *nirE* is the only one of the operon conferring SUMT activity in the pathway.

The formation of PC-2 and SHC by NirE is shown for the first time by incubation of the soluble extract with the proteins required for the synthesis of PC-2 and SHC, with NirE. Furthermore, the activity of NirE in our SHC coupled assay also indicates that NirE is performing the transformation of UIII to PC-2. The absence of NirE halts the reaction and the absence of SAM from the reaction demonstrates NirE is dependent on SAM to perform this reaction as predicted. The absence of SAM in the reaction mixture does not form SHC because the substrate for this reaction, PC-2 is not being formed. The absence of NAD<sup>+</sup> from the reaction mixture also does not produce SHC as SirC activity is NAD<sup>+</sup> dependent, but in this case we clearly have the formation of PC-2 by NirE. We can also say with confidence that the origin of the methyl groups at C2 and C7 is S-adenosyl-L-methionine, and that the reaction is likely to proceed via precorrin-1 as in the case of CobA in the vitamin B<sub>12</sub> pathway and other SUMT enzymes.

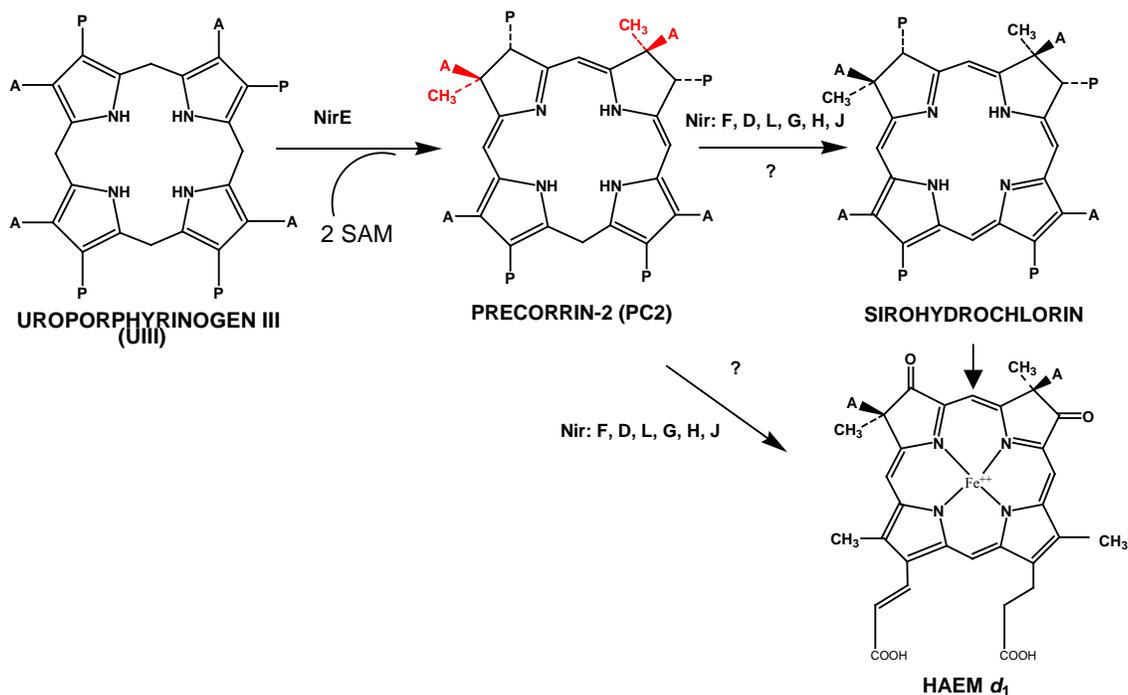
Therefore, in summary we have shown for the first time that NirE is indeed a functionally active S-adenosyl-L-methionine UIII methyltransferase, transforming UIII to PC-2. As predicted, chromatographic analysis of NirE by gel filtration

proves it is a homodimer with a subunit mass of 29 kDa, consistent with the topology of CobA, CysG and CbiF. It portrays all the defined characteristics that are attributed to similar SUMTs such as the accumulation of porphyrinoid material. It is SAM dependent, and able to generate PC-2 complementing the pathways for vitamin B<sub>12</sub> and sirohaem, providing the substrate for the generation of SHC. This suggests that SHC is may be a part of the pathway to haem *d*<sub>1</sub>, but as yet no candidate from the *nir* operon has been identified for this dehydrogenation reaction and warrants further investigation.

Thus, the first step in the haem *d*<sub>1</sub> pathway is the consecutive SAM dependent methylation of C2 and C7 from UIII to PC-2. The remaining steps performed by NirF, NirD, NirL, NirG, NirH, and NirJ to haem *d*<sub>1</sub> are as yet unknown (Figure 3.20).

### Figure 3.20: The first step of haem *d*<sub>1</sub> biosynthesis

In this chapter, NirE is confirmed to catalyse the first step in the methylation of UIII at C2 and C7 to form PC-2. It is unclear if the pathway then proceeds via SHC or from PC-2 to haem *d*<sub>1</sub>, with NirF, NirD, NirL, NirG, NirH, and NirJ.



**Optimisation of Production and Stabilisation  
of Proteins Involved in Haem  $d_1$  Synthesis**

## 4. Optimisation of production and stabilisation of proteins involved in haem $d_1$ synthesis

### 4.1. Introduction

Haem  $d_1$  is a unique tetrapyrrole found only in bacterial dissimilatory nitrite reductase, acting as the active centre for nitrite reduction. Whilst there exists a lot of literature on the enzyme for which it is a co-factor, the biosynthesis of haem  $d_1$  still remains poorly understood. This chapter looks at the overproduced recombinant proteins implicated in haem  $d_1$  biosynthesis in both *E. coli* and in *Ps. aeruginosa*, in an effort to understand the roles played by the individual components of the pathway. The previous chapter details NirE and the confirmation of SUMT activity in the transformation of UIII to PC-2. The remaining genes were cloned into appropriate vectors to allow their encoded products to be overproduced in both organisms. This study highlights the difficulties in overproduction of these proteins and furthermore the challenges incurred in stabilisation of these proteins, as well as the formulation of a new homologous expression system in *Ps. aeruginosa*.

Mutation and complementation analysis of the *nir* genes necessitated a requirement of *nirE*, *nirJ*, *nirH*, *nirG*, *nirL*, *nirD* and *nirF* for haem  $d_1$  biogenesis (Kawasaki, Arai et al. 1995; Palmedo, Seither et al. 1995; Kawasaki, Arai et al. 1997). These are well known to exist as part of operons close to the genes required for the assembly of cytochrome  $cd_1$ , and additionally form part of the denitrification gene cluster implicating their role in denitrification, as detailed in Chapter 1.

In *Ps. aeruginosa*, *nirCFDLGHJEN* are found in one locus downstream of *nirS* and are hypothesised to encode the entire set of proteins required for the formation of haem  $d_1$  (Figure 4.1) (Kawasaki, Arai et al. 1995; Kawasaki, Arai et al. 1997). In *Ps. stutzeri* the same genes for haem  $d_1$  biosynthesis are located over two loci, *nirJEN* and *nirCFDLGH*, which lie upstream and downstream of the structural gene *nirS*, respectively (Palmedo, Seither et al. 1995).

In *Pa. denitrificans*, the presence of *nirECFD* downstream of *nirS* (the structural gene for cytochrome  $cd_1$ ), has been well documented (de Boer, Reijnders et al. 1994). In this study, our analysis of the sequence downstream of this (courtesy of Prof Ferguson, Oxford) has identified a cluster containing the gene

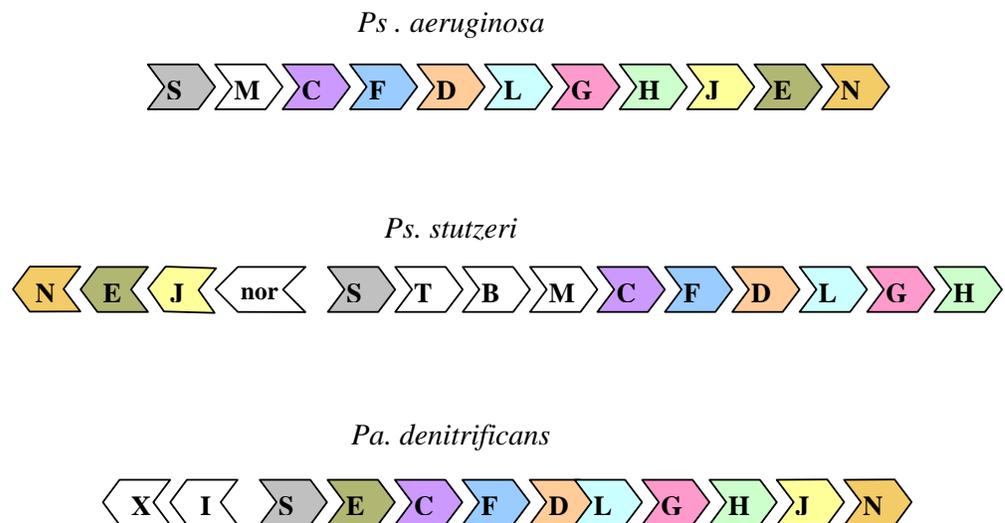
order of *nirECFDLGHJN*. It is proposed that NirF of *Pa. denitrificans*, is a fusion of NirF and NirD, the first 350 residues showing 53.1 % identity to NirF from *Ps. stutzeri* and the latter showing 50.6% identity to NirD from *Ps. stutzeri* (de Boer, Reijnders et al. 1994; Palmedo, Seither et al. 1995). However, our analysis of the *Pa. denitrificans nir* operon demonstrates this to be wrong. Whilst, the NirF protein sequence shown in the paper starts correctly, the latter 20 residues are incorrect and are derived from an erroneous frameshift. This frameshift is in the same reading frame as NirD, thus, it was presumed that NirF is a fusion of NirF and NirD because a termination of NirF was not detected. Consequently this assumption has no basis and may be ignored. Nevertheless, further analysis in this study, revealed that the separate open reading frames for *nirD* and *nirL* observed in *Ps. aeruginosa* and *Ps. stutzeri* are fused into a single open reading frame in *Pa. denitrificans*. Therefore, in *Pa. denitrificans* NirD and NirL are fused into a single protein. *Pa. denitrificans* is unique in comparison as it makes so much nitrite reductase that the periplasm preparation itself is brown. This is particularly useful as the level of haem  $d_1$  synthesis and the required enzymes may also be high.

The presence of a haem  $d_1$  like operon has recently been proposed in heliobacteria (*Heliobacillus mobilis* and *Heliophilium fasciatum*) by bioinformatic analysis (Xiong, Bauer et al. 2007). This is said to contain NirD, NirL, two forms of NirJ (NirJ1 and NirJ2) and CysG<sup>A</sup> and CysG<sup>B</sup> as well as the genes for the production of UIII from ALA. It has been hypothesised that CysG<sup>A</sup> acts as a SUMT to yield PC-2 and CysG<sup>B</sup> yields SHC, and that only NirJ is involved in decarboxylation and then CysG<sup>B</sup> is involved further in ferrochelation in the synthesis of haem  $d_1$ . The validity of this hypothesis is debatable as not all of the transformations from UIII to haem  $d_1$  are explained. It does not take into account the formation of an acrylate at C17 nor the conversion of the propionates to oxo groups at C3 and C8. NirD and NirL are proposed to not be involved in the transformation of UIII to haem  $d_1$  but only in transcriptional regulation and DNA binding. As already mentioned, *nirF*, *D*, *L*, *G*, *H*, *J*, *E*, are essential to the pathway, thus if the genes for haem  $d_1$  were present as hypothesised by Xiong (NirJ1 and NirJ2; CysG<sup>A</sup> and CysG<sup>B</sup>), the insertional mutations of *nirD* and *nirL* would have no effect on haem  $d_1$  synthesis. It is very likely that the presence of these genes are for that of an alternative pathway for haem previously observed in *D. vulgaris*, whereby UIII is methylated instead of decarboxylated (Ishida, Yu et al. 1998). This pathway has also been identified in *M.*

*barkeri*, where a cluster of genes with the enzymes encoding for UIII (*hemA*, *B*, *C*, *D*) and genes with similarity to *nirE*, *D*, and *J* have been found, but the late haem genes are missing. Methylation occurs as part of the haem pathway therefore it is presumed that the synthesis of haem employs similar enzymes with as the Nir pathway (Buchenau, Kahnt et al. 2006). This is likely to be the case in this pathway observed in Heliobacteria. Therefore this theory has no basis in the  $d_1$  pathway as they are based on bioinformatic analysis and manipulation only, they are not experimentally derived. Moreover, heliobacteria contain bacteriochlorophyll and have no indication of a requiring an alternative respiratory nitrogen metabolism pathway such as one requiring haem  $d_1$ .

**Figure 4.1: The haem  $d_1$  operon.**

Illustration of the *nir* operon from *Ps. aeruginosa*, *Ps. stutzeri* and *Pa. denitrificans*. Block arrows show the direction of transcription and homologues are shown in the same colour. The genes have not been drawn to scale. Nir X and I are shown to be involved in transcriptional regulation of the operon. This study reveals that *Pa. denitrificans* has an operon similar to that observed in *Ps. aeruginosa* except that *nirE* resides at the start of the operon preceding *nirC* as opposed to the end of the operon. Additionally, *nirD* and *nirL* are observed to be fused *Pa. denitrificans*.



## RESULTS AND DISCUSSION.

### 4.2. Approaches employed in the analysis of NirF, NirD, NirL, NirG, NirH and NirJ

NirF, D, L, G, H, J were subject to various methods of investigation to aid in the understanding of their role in the haem  $d_1$  pathway. Bioinformatic analysis highlighted the ambiguity that exists for each of the Nir proteins and thus the need for investigation of protein structure and function. This would enable an understanding of the reactions they may catalyse in the synthesis of haem  $d_1$ . The different methods employed are discussed here first before examination of the results obtained

#### 4.2.1. Analysis of rare codons within the *nir* genes

The level of rare codon usage was checked for *nirF*, *nirD*, *nirL*, *nirG*, *nirH*, and *nirJ* to eliminate any problems in the expression of these genes. Each organism carries its own bias in the usage of the 61 available codons. When the mRNA of heterologous genes are overexpressed in *E. coli*, differences in codon usage can impede translation due to the demand of one or more tRNA that maybe rare or lacking. Examination of codon usage in *E. coli* reveals the underrepresentation of AGA, AGG, CGA, CGG, GGA, AUA, CUA and CCC.

The number of these rare codons in the genes were calculated and found to be minimal and are summarised in Table 4.1. Therefore, expression of these genes within a host carrying these rare codons is not required, as it would have no implication on protein production.

**Table 4.1: Tabulation of rare codon usage for each *nir* gene**

Table showing percentage of rare codons calculated for each gene, by checking for AGA, AGG, CGA, CGG, GGA, AUA, CUA and CCC rare codons in *E. coli*, and represented as a percentage of the total codons. The level of rare codons is low and the use of a host with these codons would not affect protein production.

Gene	<i>nirF</i>	<i>nirD</i>	<i>nirL</i>	<i>nirG</i>	<i>nirH</i>	<i>nirJ</i>
Rare codon usage	3 %	2.6 %	2.3 %	0 %	1.2 %	3.6 %

#### 4.2.2. Cloning and heterologous overproduction of proteins in *E. coli*

PCR primers were designed for each of the genes *nirF*, *nirD*, *nirL*, *nirG*, *nirH*, and *nirJ* as detailed in Chapter 2, Table 2.8. All *nir* genes were amplified by PCR using the respective designed primers. These were constructed to introduce suitable restriction sites, *NdeI* and *BamHI*, at the 5' and 3' termini of each *nir* gene respectively. Introduction of these sites allows in frame cloning into the pET14b ( $\text{amp}^R$ ) expression vector (via pGEMT-Easy) permitting the encoded protein to be produced as a N-terminal poly-Histidine tagged fusion protein, thereby facilitating rapid purification by metal chelate chromatography.

To facilitate protein overproduction, the pET14b-*nir* gene plasmid constructs were individually transformed into *E. coli* BL21 pLysS (DE3,  $\text{chlor}^R$ ) strains in which the T7 polymerase is induced by the addition of IPTG, to a final concentration of 0.4 mM in cultures grown in LB medium at 37 °C (plus ampicillin 100 mg/l and Chloramphenicol 34 mg/l) with aeration. Cultures were further incubated as appropriate at 37 °C or 16 °C with aeration, before harvesting by centrifugation (10 000 x g, 4 °C). The bacterial pellets were resuspended in binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 8.0). Proteins were purified as described in the pET system manual (Novagen). Typically, prior to purification, the crude cell fractions were sonicated and centrifuged (10 000 x g, 4 °C). The resultant supernatants containing soluble protein were then loaded onto separate charged His-Bind columns. The columns were washed with binding buffer, and wash buffers of increasing imidazole concentration (50 mM then 100 mM), before eluting the protein in elution buffer (400 mM imidazole). The presence of protein in the fractions was detected by the Biorad protein assay and SDS-PAGE analysis.

##### 4.2.2.1. Protease inhibitors

Primary analysis of overproduction in *E. coli* showed the Nir Proteins were quite unstable and prone to degradation. As a result all subsequent purifications were performed in the presence of EDTA free protease inhibitors added to the cell fraction prior to sonication. This allowed for protein integrity of the Nir proteins from enzymes and proteases present within the cell. The effect of including protease inhibitors is most evident for NirF, shown in Figure 4.16B (p.160).

### 4.2.3. Control of expression by control of [IPTG] within a host strain

Since, the overproduction was observed to be very high in all the Nir proteins, but the amount of soluble protein low, this implied that the level of overproduction within the cell might be too high to withstand. Thus, controlling the level of overproduction may in turn increase the solubility of the encoded proteins. This is facilitated by controlling the concentration of IPTG, which in turn requires transformation into a cell line that allows controlled levels of expression, unlike BL21 pLysS (DE3 chlor<sup>R</sup>), which has maximal expression levels.

Tuner pLaci (chlor<sup>R</sup>) enables adjustable levels of protein expression throughout all cells in a culture. They are *lacYZ* deletion mutants of BL21, where the *lac* permease (*lacY*) deletion allows a constant level of entry of IPTG into all the cells, producing a concentration-dependent and homogeneous level of induction. In this manner the final concentration of IPTG induction was decreased four-fold to 0.1 mM and ten-fold to 0.04 mM, from 0.4 mM. The resultant samples for each NirF, NirD, NirL, NirG, NirH, and NirJ at both levels of decreased induction were purified and analysed by SDS-PAGE.

### 4.2.4. Cloning & homologous overproduction of proteins in *Ps. aeruginosa*

A new strategy was employed to overcome the lack of solubility and stability of the proteins, as observed in *E. coli*, by the overproduction of the recombinant proteins in their native organism, *Ps. aeruginosa*. Since *Ps. aeruginosa* does not recognise the T7 promoter, a new system was formulated to facilitate recombinant protein overproduction in *Ps. aeruginosa*. Various plasmids were screened for transformation efficiency in *Ps. aeruginosa* and for suitable multiple cloning sites. The plasmid most compatible in both cases was pUCP19-Nco / Nde a *Escherichia-Pseudomonas* shuttle vector for recombinant protein overexpression in *Pseudomonas* (Cronin and McIntire 1999).

All the genes *nirF*, *nirD*, *nirL*, *nirG*, *nirH*, and *nirJ* genes were successfully individually subcloned into pUCP19-NcoI (carb<sup>R</sup>), from pET14b using the restriction sites *NcoI* and *BamHI*. This allowed both the gene and His-tag to be subcloned into pUCP19-NcoI and allowing protein overproduction and purification by metal chelate chromatography.

To facilitate protein overproduction, plasmid constructs were individually transformed into wild type *Ps. aeruginosa* (PAO1), in which the *lac* promoter is induced by the addition of IPTG, to a final concentration of 1 mM in cultures grown in LB medium at 37 °C (plus carb 500 mg/l) with aeration. Cultures were further incubated at room temperature with aeration, before harvesting by centrifugation (10 000 x g, 4 °C). Anaerobic cultures were grown in the same way without aeration and on a nitrate-containing medium (25 mM) with mineral oil to deplete oxygen. The bacterial pellets were resuspended in binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 8.0). The resultant homologously expressed recombinant proteins were then purified by metal chelate chromatography. The presence of protein in the fractions was detected by the BioRad protein assay and protein integrity checked by SDS-PAGE or western blotting.

#### **4.2.5. Protein stabilisation screen**

In an effort to stabilise the Nir proteins to obtain a sufficient level for further characterisation, various buffers were screened for each protein. All previous investigations were performed in 50 mM Tris pH8, 0.5M NaCl, but the proteins were observed to be unstable. The pI was also taken into account for each protein (see appendix). The buffers used in the stabilisation screen included 50 mM Tris pH8.0; 50 mM CHES pH10, 9.5, 9.0; 50 mM MES pH6.5, 6.0; 50 mM Hepes pH8.5, 8.0, 7.5; 50 mM sodium phosphate pH 8.5, 8.0; 50 mM potassium phosphate pH8.5, 8.0, each with 0.5M NaCl. The same buffers were screened again at 50 mM concentration but at a salt concentration of 0.1M, then each buffer at 20 mM concentration with 0.5M or 0.1M salt.

Primarily, proteins from 4 x 1L growths for each protein were purified using ice-cold buffers and eluted on ice in 50 mM Tris pH8, 0.5M NaCl. The fractions containing the desired protein were then pooled and buffer exchange was performed on a pD10 column eluting the protein into the buffer being tested. This was performed in batches of 6 of the same protein in different buffers. Firstly, screening was performed for the following buffers: 50 mM Tris pH 8.0, 50 mM CHES pH 9.5, 50 mM MES pH 6.5, 50 mM Hepes pH 8.0, 50 mM sodium phosphate pH 8.5, 50 mM potassium phosphate pH 8.0 each with 0.5M NaCl. The protein solutions were each then concentrated in 2ml vivaspin columns. The protein was deemed stable in a buffer if one was able to sufficiently concentrate it without the protein

precipitating out of solution. If a protein was found to be more stable in a particular buffer, then the buffer was further used to screen for further stability at different pH or NaCl concentration if applicable. If no stabilisation was observed the same buffers at 50 mM were used to screen for stability again but at a lower 0.1M salt concentration and further at a lower, 20mM buffer concentration and varying NaCl concentrations of 0.5 mM and 0.1 mM.

Each of these approaches was applied to NirF, NirD, NirL, NirG, NirH, and NirJ. The results obtained for each protein are discussed next.

### 4.3. NirD, NirL, NirG, NirH

#### 4.3.1. Bioinformatic analysis

The functions of NirD, NirL, NirG and NirH are to date unknown, as is their structure and any reactions they may catalyse in the haem  $d_1$  pathway. Comparison of these proteins to those of known structure and function provides very limited information. A similarity of NirD, NirL, NirG and NirH is observed only to a family of leucine responsive protein (LRP) transcriptional regulators (Figure 4.2). Therefore the functions of NirD, L, G, H are somewhat ambiguous and remain to be characterised.

Indeed, high levels of identity between NirD and NirG (Figure 4.3A, 41 %) and NirL and NirH (Figure 4.3B, 33 %) may suggest the catalysis of symmetrical reactions, such as decarboxylation of C12 and C18 or elimination of propionic side chains or oxidation at C3 and C8. Furthermore, they may form a multimeric and multifunctional enzyme complex in the form of a  $\alpha\beta\alpha'\beta'$  unit, as they are proposed in the *nirD* locus of *Ps. stutzeri* (*nirDLGH*) (Palmedo, Seither et al. 1995). The similarity of NirD and NirG, which are composed of 150 (17 kDa) and 147 (16.5 kDa) amino acids respectively, and of NirH and NirL being 171 (19.6 kDa) and 174 (20.2 kDa) respectively, also supports this view. The degree of identity that exists within all four proteins (Figure 4.3C), suggests that these may have arisen as result of gene duplication events. Identity between NirD, NirL, NirG and NirH of *Ps. aeruginosa* with those from *Ps. stutzeri* is in the region of 60 % to 72 % and 39 % to 62 % for *Pa. denitrifications* (Table 4.2). This indicates a high degree of conservation of these proteins amongst species and may be indicative of the evolution of a highly specialised role. In *heliobacteria* only NirD and NirL are found and proposed by the author to regulate the transcription expression of *nir* genes only (Xiong, Bauer et al. 2007). However, it is also well documented that FNR-like transcriptional regulators mediate transcription regulation of the *nir* operon. Consequently no clear function can be assigned to these proteins and there remains ambiguity in their roles as part of the haem  $d_1$  pathway.



**Figure 4.3: Sequence alignments of NirD, NirL, NirG, NirH.**

Alignments of NirD with NirG from *Ps. aeruginosa* (PAO) shows 47 % identity (A) and NirH with NirL from *Ps. aeruginosa* shows 33 % identity (B). Alignment of NirD, NirL, NirG, and NirH (C) shows the degree of identity that exists between them, suggesting that these may have arisen as result of gene duplication events. The asterisks show identical amino acids, dots those belonging to the same group.

(A)

NirD	MDDL	SRRL	LARY	QKGL	PICA	EPYR	MAET	LCSE	AEVLE	RRLR	LEAD	GALSR	VG	PVLR	HQ	60			
NirG	MDEF	DRLL	LNRL	QHGL	PLEPH	PYALL	AAEL	DCRE	EILR	RLDD	LLDD	GGTL	TRFG	PLFD	IE	60			
	***:	.****	* *	*****:	..**	:*	*.*	* *:	*.**	* *	***:	*.***:	:	:	:				
NirD	RAG-	ASTL	AALAV	PEER	LQ	RV	AERIS	QYAE	VNHN	YQ	REH	RYNL	WF	VL	TAGD	RAQL	DRVLA	119	
NirG	PLGG	AFTL	AAMSV	PEAR	FEEI	AA	LLAG	WP	QVAH	NYR	RE	HALN	MW	LV	ACD	SPA	EVAET	LA	120
	* *	*****:	***	*:.*	:	:	..**	***:***	**:	*:..*	:	:	:	:	:	:	:	:	
NirD	EIAA	DTGL	QPLD	LP	ME	AY	CIDL	AF	PLE	ASR	150								
NirG	RLER	ESGL	AVLD	LP	KE	ATYH	VGL	HFP	L---	147									
	..	:	***	****	:	:	*.*	***	:	:									

(B)

NirL	-MNP	VEPL	AAPQ	RQH	RYL	LE	QGLP	LASR	PYR	VLA	ERIG	AGED	EVLE	QVRR	WDED	GLFRR	59				
NirH	MSAC	ISPD	DALL	ARRL	IELT	QAGL	PLVAD	PWAW	IAAQ	LR	LSE	AETL	LALL	KRLR	DAGV	IRR	60				
	.*	*	:	*	:	*	:	****:	.*	:	.*	:	.*	:	.*	:	****				
NirL	FGV	ILH	HRAL	GYT	ANAM	LVLD	VADAE	VD	AVGR	ALAH	ETIV	SLCY	RRPR	RLP	MW	PYNL	FCM	119			
NirH	IAAV	PNH	YRLG	YRH	NGMT	VWDV	ADERI	ERL	GRLV	GG	LSFV	SHCY	RRPR	HLP	QW	RYNL	FAM	120			
	..:	.*	***	*.*	*	****	..:	.*	..	:	***	*****:	* *	****	.*						
NirL	IHG	RE	GE	VE	RQ	IE	ALL	ER	HALR	QT	PHR	WFL	SL	RAYK	QCG	GRY	TAPP	ADL	ERR	HG	174
NirH	VHGR	SEAE	IEG	YR	QQI	RLL	LGED	CR	ADE	MLV	SSR	ILK	KTG	---	LRL	AQ	KE	ER	PC	171	
	***:	..**:	:	:	:	:	..	..	..	..	..	..	..	..	..	..	..	..	..	..	

(C)

NirL	-MNP	VEPL	AAPQ	RQH	RYL	LE	QGLP	LASR	PYR	VLA	ERIG	AGED	EVLE	QVRR	WDED	GLFRR	59				
NirH	MSAC	ISPD	DALL	ARRL	IELT	QAGL	PLVAD	PWAW	IAAQ	LR	LSE	AETL	LALL	KRLR	DAGV	IRR	60				
NirD	-----	MDDL	SRRL	LARY	QKGL	PICA	EPYR	MAET	LCSE	AEVLE	RRLR	LEAD	GALS	R	52						
NirG	-----	MDEF	DRLL	LNRL	QHGL	PLEPH	PYALL	AAEL	DCRE	EILR	RLDD	LLDD	GGTL	TR	52						
		***:	..**	:	***:	..**	:	***:	..**	:	***:	..**	:	***:	..**	:					
NirL	FGV	ILH	HRAL	GYT	ANAM	LVLD	VADAE	VD	AVGR	ALAH	ETIV	SLCY	RRPR	RLP	MW	PYNL	FCM	119			
NirH	IAAV	PNH	YRLG	YRH	NGMT	VWDV	ADERI	ERL	GRLV	GG	LSFV	SHCY	RRPR	HLP	QW	RYNL	FAM	120			
NirD	VG	PVLR	HQR	AG--	ASTL	AALAV	PEER	LQ	RV	AERIS	QYAE	VNHN	YQR----	EH	RYNL	WFV	105				
NirG	FG	PLFD	IE	PLGG--	AFTL	AAMSV	PEAR	FEEI	AA	LLAG	WP	QVAH	NYRR----	EH	ALNM	WL	106				
	..	:	*	:	..	..	..	..	..	..	..	..	..	..	..	..	..				
NirL	IHG	RE	GE	VE	RQ	IE	ALL	ER	HALR	QT	PHR	WFL	SL	RAYK	QCG	GRY	TAPP	ADL	ERR	HG	174
NirH	VHGR	SEAE	IEG	YR	QQI	RLL	LGED	CR	ADE	MLV	SSR	ILK	KTG	---	LRL	AQ	KE	ER	PC	171	
NirD	LT	AGD	RAQL	DRVLA	EIAA	DTGL	QPLD	LP	ME	AY	CIDL	AF	-----	LE	ASR-	150					
NirG	VACD	SPA	EVAET	LAR	LER	ESGL	AVLD	LP	KE	ATYH	VGL	HFP	-----	L-----	147						
	:	..**	:	:	:	:	:	:	:	:	:	:	:	:	:	:					

### 4.3.2. Optimisation of overproduction

#### 4.3.2.1. Overproduction in *E. coli*

The genes *nirD*, *nirL*, *nirG* and *nirH*, from *Ps. aeruginosa*, were successfully cloned into the pET14b expression vector without mutations, behind a polyhistidine-encoding region. The encoded proteins were therefore produced as N-terminal His-tagged fusion proteins in *E. coli*, and purified by metal chelate chromatography.

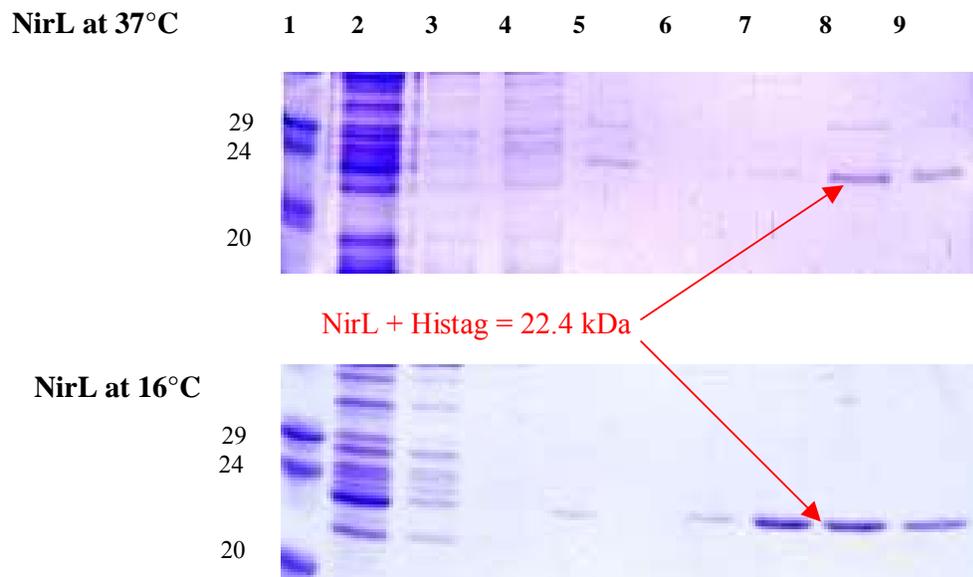
All NirD, NirL, NirG, and NirH, were detected on SDS-PAGE and shown to be purified after elution at 400 mM imidazole from the His-bind column. Notably, NirD is overproduced and purified together with an unknown protein, presumably a host encoded chaperone protein (Figure 4.7). Therefore, all the proteins whilst only being soluble to a certain degree, required screening for optimisation of conditions for maximal expression.

The first attempt towards increasing solubility was to change the conditions during induction to 16 °C, instead of 37 °C. The length of induction was also varied from 3 hours to overnight. A greater degree of solubility is observed for NirD, NirL, NirG, and NirH with induction of protein overproduction at 16 °C overnight as opposed to induction at 16 °C for 3 hours or 37 °C. The difference in yield is summarised in Table 4.3. The SDS-PAGE results are shown for NirL (Figure 4.4) and NirD (Figure 4.7), are representative of those obtained for NirG and NirH.

Induction at 16 °C overnight gives a sufficient yield of protein if performed on a large scale. However, it was not possible to sufficiently concentrate the eluted proteins because of their instability in solution. Analytical gel filtration of NirL, NirG, and NirH without concentrating them revealed that they were unstable and could not be sufficiently purified to homogeneity being present mostly as aggregates (Figure 4.5), even in the presence of protease inhibitors. NirD (Figure 4.8) could not be separated from the co-eluted protein indicating a very tight complex, discussed later. Since NirD, NirL, NirG, and NirH all behave in a similar fashion, it suggests that perhaps these proteins need to be coproduced as a multimeric and thus multifunctional complex of NirDLGH as has been reported for *Ps. stutzeri* (Palmedo, Seither et al. 1995). They may then only be sufficiently soluble and stable for further investigation.

**Figure 4.4: SDS-PAGE gels of incubation at 37 °C and 16 °C for NirL.**

NirL is shown to represent the difference in solubility revealed when cultures were induced for overproduction at the two temperatures. Similar results were obtained for NirD, NirH and NirG. The first lane is the protein marker. Lane 1 the supernatant. Lanes 2, 3, 4, 5 are washes with binding buffer of increasing imidazole concentration (0 mM, 5 mM, 50 mM and 100 mM respectively). Lanes 6, 7, 8, 9 represent the purified protein eluted with 400 mM imidazole. These results reveal the greater solubility of the protein NirL when induced at the lower temperature, 16 °C overnight.



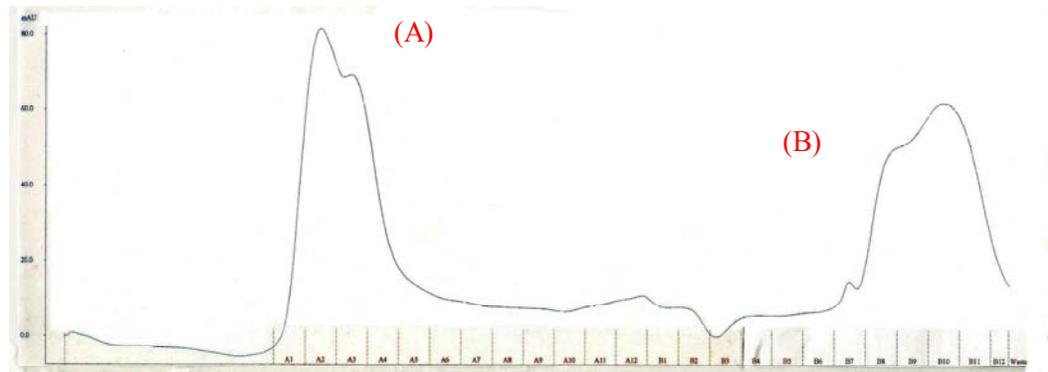
**Table 4.3: Comparison of soluble protein yields from NirD, NirL, NirG and NirH from 1L cultures incubated at 37 °C and 16 °C.**

Temp	NirD	NirL	NirG	NirH
37 °C	8.47 mg	10.95 mg	8.74 mg	9.85 mg
16 °C	15.65 mg	12.67 mg	13.01 mg	11.56 mg

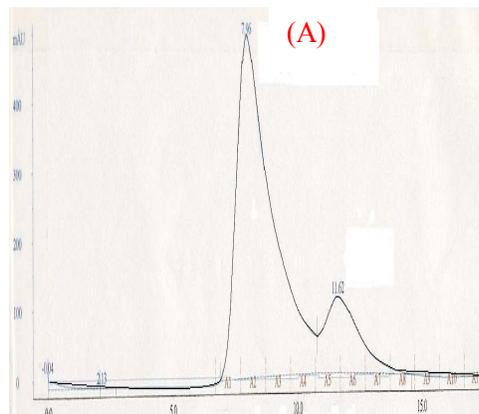
**Figure 4.5: Gel filtration elution profiles of NirL, NirG and NirH**

NirL, NirG and NirH were analysed by gel filtration chromatography, without concentrating the sample eluted from metal chelate column. The elution profiles shown indicate that NirL, NirG and NirH are unstable the majority being present as aggregated protein (A) or broken peptides (B)

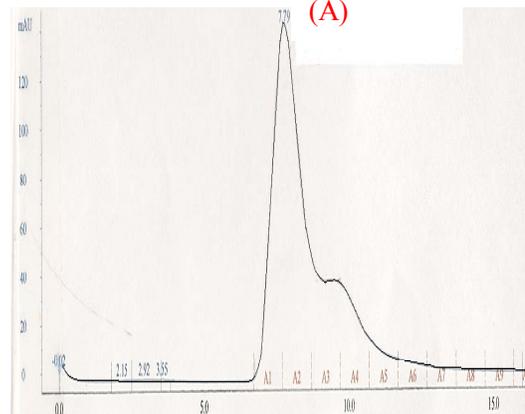
**NirL**



**Nir G**



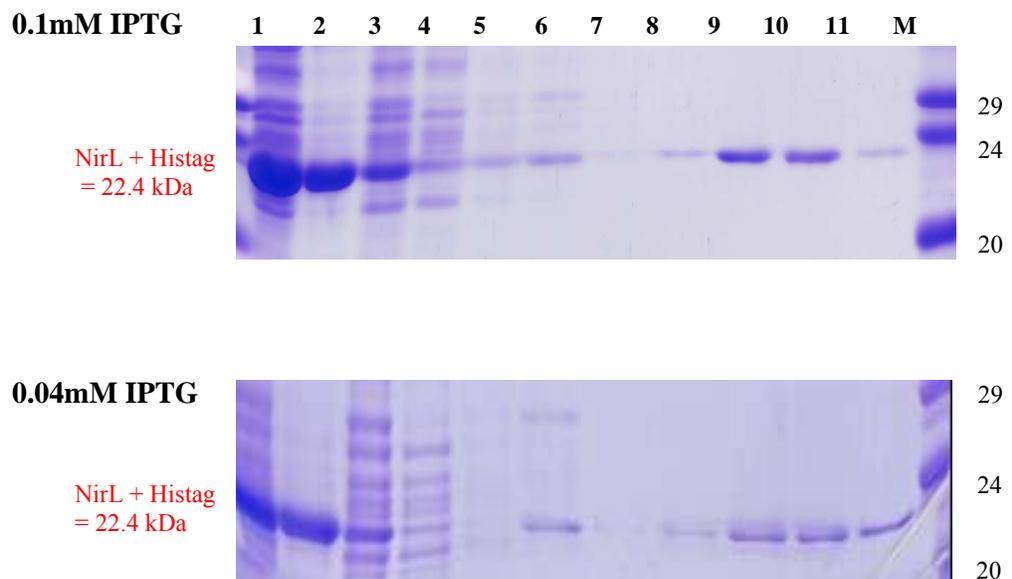
**NirH**



The majority of the protein was observed to be present in the cell debris fraction after sonication. Therefore, the level of overproduction was high but the amount of soluble protein was low in comparison. Thus, adjusting the levels of protein expression by using Tuner cells may in turn increase the amount of soluble protein, such that they are less likely to be drawn in to inclusion bodies. The final concentration of IPTG induction was decreased four-fold to 0.1 mM and ten-fold to 0.04 mM, from 0.4 mM, and cell fractions analysed by SDS-PAGE. The results showed no change in the yield of soluble NirD, NirL, NirG and NirH at either decreased concentration of IPTG, even in the event of reduced overproduction (Figure 4.6).

**Figure 4.6: SDS-PAGE of NirL at 0.1mM IPTG 0.04mM IPTG in Tuner Cells**

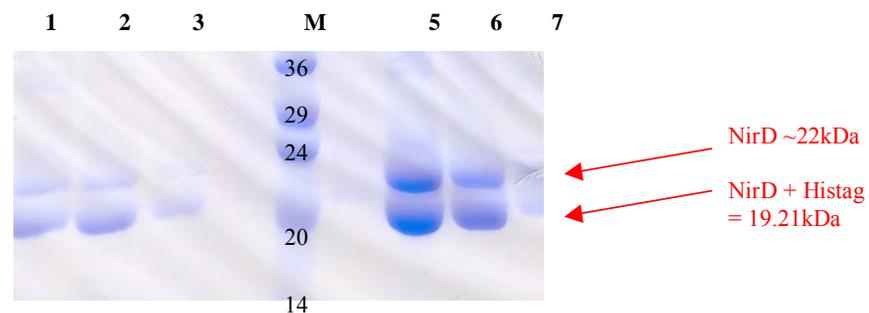
NirL is representative of the results obtained for NirD, NirG and NirH. A decrease in the level of expression is observed with a decrease in [IPTG], seen in lanes 1, but no increase in solubility is observed for NirG, NirH or NirL, at either levels of IPTG induction. M represents the protein marker, 1 is the total crude cell extract, 2 is the pellet after lysis, 3 the supernatant, lanes 4, 5, 6, and 7 are washes with binding buffer of increasing imidazole concentration (0 mM, 5 mM, 50 mM and 100 mM respectively), lanes 8, 9, 10, 11 are the 2 ml protein fractions eluted with 400 mM imidazole.



NirD purification by metal chelate chromatography revealed that NirD has great affinity to a protein of size 22 kDa. A change in the conditions and incubation did not alter the level of this co-expressed protein as seen in Figure 4.7. As with NirL, NirG and NirH, NirD incubation at 16 °C gives a greater amount of soluble protein. Purification was repeated in the presence of protease inhibitors, and concentrations of imidazole in the washes prior to elution were changed, but this extra protein remained firmly associated with NirD. Furthermore, they cannot be separated from each other even by FPLC as is evident elution profile and the SDS-PAGE gel of these fractions in Figure 4.8. This indicates that they are tightly bound together with the concentration of both being proportional to each other, suggesting a tight complex. Further analysis of this co-expressed protein by MALDI revealed the protein being co-expressed is NirD itself from *Ps. aeruginosa*, having on average a 62 % - 87 % match. It is interesting that this protein is not the size of NirD that is 19kDa with a His-tag, and is approximately 3 kDa larger at 22 kDa, thus we do not have a typical homodimer. The protein at 22 kDa may be suggestive of some form of post-translational modification of NirD, and it is notable that it remains very tightly bound to the original NirD at 19 kDa. It is also suggestive of the preference of these proteins to co-exist as larger molecules and indeed as the NirDLGH complex previously hypothesised. Consequently, co-crystallation of this tightly bound complex of NirD would be invaluable in identification of structure, stability and function.

**Figure 4.7 SDS-PAGE gel of NirD incubated at 37 °C and 16 °C**

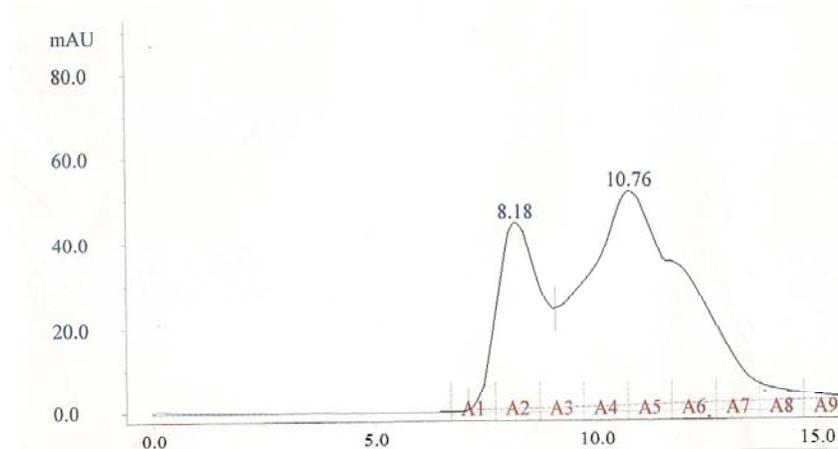
NirD is clearly overproduced at 19.21 kDa with a host encoded protein that is ~22 kDa. Lanes 1, 2, 3, represent purified NirD eluted with 400 mM imidazole from cultures incubated at 37 °C. Lane M is protein marker. Lanes 5, 6, 7 are of purified NirD incubated at 16 °C. Clearly the yield of soluble protein is greater at 16 °C.



**Figure 4.8: NirD gel filtration elution profile and SDS-PAGE of eluted fractions.**

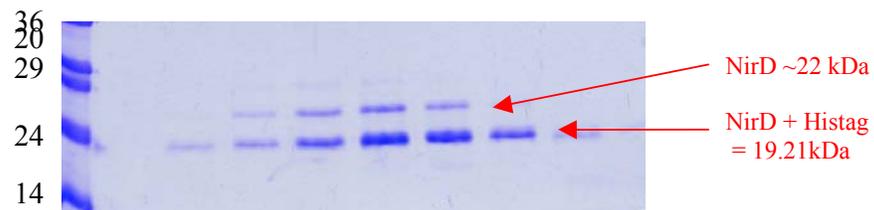
Gel filtration was not able to separate the two NirD proteins of different size at 19 kDa and 22 kDa (A). Any other contaminants were successfully removed. The yield of the co-expressed 22 kDa NirD seems to be proportional to the yield of NirD at 19 kDa as seen in the SDS-PAGE of the eluted fractions (B). M is the protein size marker and A1 to A9 are the 1 ml fractions from the gel filtration column represented in the elution profile.

(A). NirD elution profile



(B). SDS-PAGE of the corresponding fractions from gel filtration.

A1 to A9 corresponds to the fractions eluted from the gel filtration column as per the elution profile above.



#### 4.3.2.2. Overproduction in *Ps. aeruginosa*

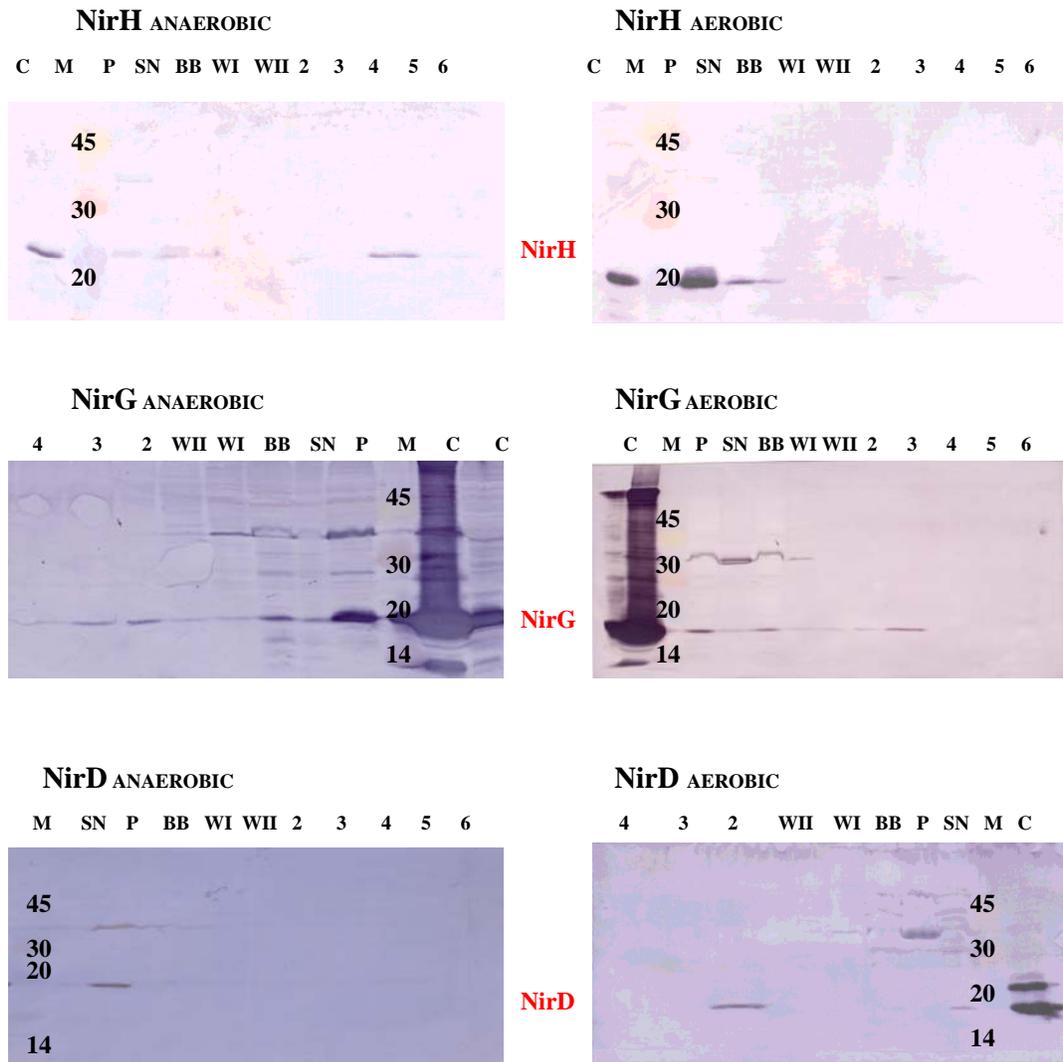
Attempts were made to overcome the lack of solubility and stability of NirD, NirL, NirG and NirH, when overproduced in *E. coli*, by the overproduction of the recombinant proteins in their native organism, *Ps. aeruginosa*. All *nirD*, *nirL*, *nirG* and *nirH*, were individually subcloned into pUCP19-NcoI, from pET14b using the restriction sites *NcoI* and *BamHI*. This allowed transfer of both the gene and the encoded His-tag into pUCP19-NcoI and thus recombinant protein overproduction in *Ps. aeruginosa* and purification by metal chelate chromatography. The presence of protein from 1L cultures of *Ps. aeruginosa* grown aerobically or anaerobically and purified was analysed by SDS-PAGE and western blotting.

The level of overproduction of recombinant NirD, NirL, NirG, NirH in *Ps. aeruginosa* was observed to be very low under both aerobic and anaerobic conditions. This was not clearly detectable on SDS-PAGE gels, thus western blotting towards the His-tag was performed. The resultant western blots for NirG and NirH, (Figure 4.9), are shown but the same result was observed for NirL and despite similar efforts as performed in *E. coli*, to change the incubation temperature, no dependency for temperature was observed. Anaerobiosis was hypothesised to increase the level of expression, as expression in the presence of nitrate and the absence of oxygen is shown to increase the expression of the *nir* genes (Filiatrault, Wagner et al. 2005). However, again no clear dependency on the presence or absence of oxygen was observed in comparison to aerobic overproduction in *Ps. aeruginosa* and aerobic overproduction in *E. coli*.

Results obtained for NirD (Figure 4.9) were particularly intriguing; no extra band was observed for over production in *Ps. aeruginosa* as oppose to the extra NirD band observed at 22 kDa in *E. coli*. Therefore the presence of the extra NirD in *E. coli* maybe due to the overproduction of this protein in an organism that does not produce this normally and may be artifactual. Furthermore, NirD overproduced in *E. coli* was used as a control and western blotting displayed 2 bands for this sample. This indicates that it too is His-tagged, in addition to NirD at 19.2 kDa that was expected. It is certainly fascinating that these two remain tightly bound but are of different size. Since protein levels in *Ps. aeruginosa* are quite low to be able to obtain a sufficient amount of protein for analysis, this returns us to the expression of the *nir* genes in *E. coli* and to the investigation of the theory that NirD, NirL, NirG, NirH may exist together as a multimeric protein.

**Figure 4.9: NirG, NirH and NirD western blots from overproduction in *Ps. aeruginosa* aerobically and anaerobically**

NirD, NirL, NirG, NirH are not clearly visible on SDS-PAGE but are seen on western blots by raising antibodies to the His-tag. NirH, NirG and NirD are shown: **C** is the control purified from *E.coli*. **M** is the protein marker. **SN** represents the supernatant and **P** the pellet after sonication. **BB, WI, WII** are washes with buffer of increasing imidazole (5 mM, 50 mM, 100 mM respectively). The numbered lanes represent purified protein eluted at 400 mM imidazole. These results reveal the lower level of overproduction in comparison to the controls, from *E. coli*. Similar results were obtained for NirL. There is no clear dependency on protein overproduction with aerobic or anaerobic growth. In some cases an extra band is seen but this band at ~40 kDa is also seen in the control (not shown). NirD = 19.21 kDa, NirH = 21.78 kDa, NirG = 18.75 kDa.



#### 4.3.2.3. Co-expression of NirD, NirL, NirG, and NirH.

The complementarity of NirD, NirL, NirG, and NirH is well observed (Figure 4.3), as is the similarity of these proteins amongst denitrifying bacteria (Table 4.2). The similar behaviour of the overproduced proteins in this study and the partiality of these proteins to form complexes is suggestive that these proteins to prefer to exist as a complex. The presence of a complex of NirDLGH as a multimeric protein has been proposed before (Palmedo, Seither et al. 1995). However, it has not been proven experimentally. Therefore, the co-expression of these proteins together would be of great significance if they indeed form a complex.

To investigate this possibility in both *E. coli* and *Ps. aeruginosa*, pET14b (pVP76) and pUCP-Nco (pVP73) were respectively constructed containing *nirD*, *nirL*, *nirG* and *nirH*, such that only *nirD* would be encoded with an N-terminal His-tag. Subsequent overexpression of this plasmid was performed in *E. coli* and *Ps. aeruginosa* as before, with subsequent purification by metal chelate chromatography. If NirD, NirL, NirG, and NirH indeed prefer to exist as a complex, purification on a His-bind column would enable NirD to also co-purify with NirL, NirG, and NirH. A higher similarity exists of NirD with NirG and it is predicted that NirD would have a greater affinity for NirG.

Primarily, a small difference is observed from NirD and NirG overproduced alone in *E. coli* (BL21 DE3 pLysS) in comparison to NirDLGH overproduced together (Figure 4.10A), by SDS-PAGE analysis of the supernatant and pellet after sonication. Controls were performed with and without pET14b in *E. coli*. NirDLGH coproduced together showed additional bands in the region of 15-18 kDa that could be NirG (16.5 kDa) but no such bands corresponding to NirL (20.21 kDa) and NirH (19.58 kDa) could be clearly identified.

Subsequent purification on a His-bind column revealed a small amount of protein is purified in comparison to NirG and NirD when overproduced on their own (Figure 4.10B). Importantly, purification of NirDLGH shows no band for NirD expected at 19.21 kDa but the 22 kDa band for NirD is seen, along with a band corresponding to the size of NirG (16.55 kDa), which is not seen when NirD is purified alone. This is notable since it suggests that NirD prefers to exist at the higher mass and further that NirD binds to NirG. This is the first time that this has been shown, and is in

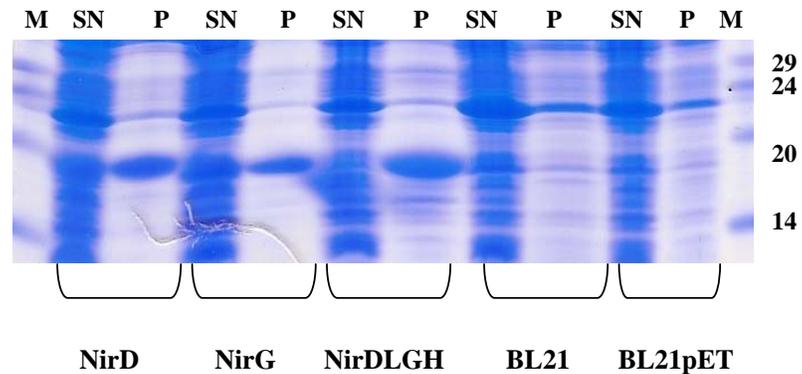
accordance with previous reports of possible complex formation and the stated complementarity of NirD to NirG. An extra band is seen at ~23 kDa but does not correspond to the size of NirL or NirH and is also observed when NirD is purified alone and thus can be eliminated.

However, SDS-PAGE analysis of NirDLGH, of the supernatant and pellet after sonication in comparison to the eluted fractions of purified protein reveals that NirD at a size of 19.21 kDa is drawn completely into the cell debris fraction (Figure 4.10C). This may be a consequence of the higher 22 kDa form of NirD being bound to NirG instead of the 19.21 kDa NirD and may be indicative of some sort of stabilisation feature of the 22 kDa form. It must also be noted that a significantly smaller amount of these proteins are observed when coproduced in comparison to NirD and NirG alone.

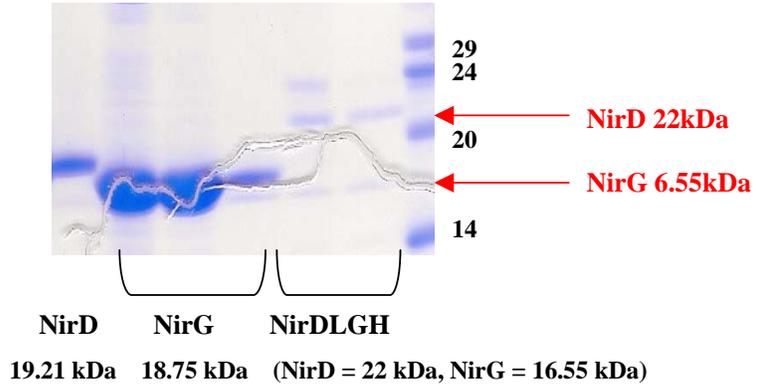
Similar results were obtained in JM109, by co expression of *nirDLGH*pUCPNco (data not shown). Co-expression in *Ps. aeruginosa* was not evident since the level of overproduction of NirD, NirL, NirG, and NirH on their own is low (data not shown). Therefore any co-production is more difficult to detect in *Ps. aeruginosa*.

**Figure 4.10: SDS-PAGE analysis of the co production of NirD, NirL, NirG and NirH in *E. coli* (BL21 DE3 pLysS)**

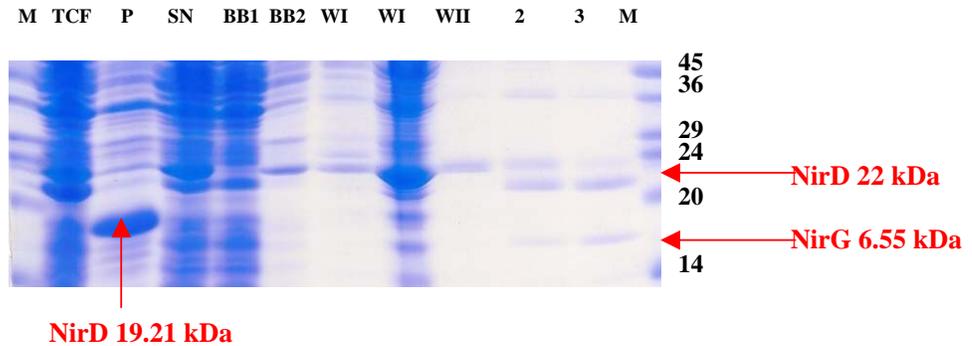
(A) SDS – PAGE analysis of the supernatant and pellet after sonication. NirD and NirG were overproduced alone and coproduced as part of NirDLGH. Controls were performed with and without pET14b in *E. coli* (BL21 DE3 pLysS). **M** is the protein marker. **SN** the supernatant and **P** is the pellet after sonication.



(B) SDS – PAGE analysis of the eluted fractions of NirD and NirG overproduced alone as well as NirDLGH overproduced together. Bands for NirD at 22 kDa and NirG at 16.55 kDa are evident in the lanes for NirDLGH co eluted at 400 mM imidazole from a His-bind column.



(C) SDS – PAGE analysis of the purification of NirDLGH. **M** is the protein marker. **TCF** is the total cell fraction. **SN** the supernatant and **P** is the pellet after sonication. **BB**, **WI**, **WII** are washes with buffer with increasing imidazole (5 mM, 50 mM, 100 mM respectively). The numbered lanes represent purified protein eluted at 400 mM imidazole where bands for NirD at 22 kDa and NirG at 16.55 kDa are evident. NirD at 19.21 kDa is seen in completely in the cell debris pellet.



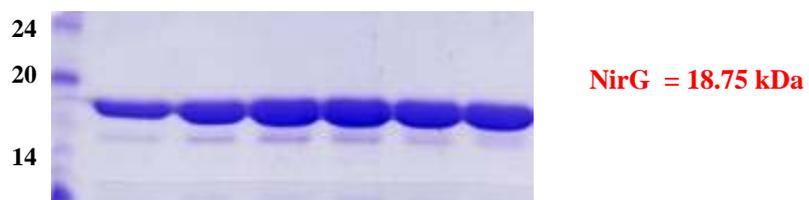
### 4.3.3. Stabilisation screen

NirD, NirL, NirG and NirH, were all screened for stability in a variety of different buffers as detailed in Section 4.2.5. Protein stability was verified by SDS-PAGE analysis and gel filtration chromatography.

NirG was the only protein found to be suitably stable in 50 mM Hepes, 0.1M NaCl. Whilst being purified in 50 mM Tris pH8, 0.5M NaCl, the average yield from 1L was 13 mg, however in Hepes a yield of 26 mg was achieved by suspension of the cell fraction and purification in Hepes. A marked difference in stability is observed both by SDS-PAGE and gel filtration analysis (Figure 4.11 and 4.12). Gel filtration chromatography further allowed the isolation of a tetramer (75 kDa) and a monomer of NirG (18.75 kDa) confirmed by native gel analysis (Figure 4.13A). The same native gel is run with NirG in the presence and absence of  $\beta$ -mercaptoethanol (Figure 4.13B), revealing that the tetramer is held together with disulphide bridges. Furthermore, observation of the SDS-PAGE gels always showed a band below that of NirG, which was predicted to be NirG without the His-tag. This band appears to increase quantity the older NirG is. Consequently, the His-tag was cleaved from NirG and run alongside a two-week-old sample of NirG, showing it is exactly the same size as cleaved NirG (data not shown). This was also the same size as the band obtained for NirG, when *NirDLGH* are co-expressed. It can be inferred that NirG is somewhat cleaving the His-tag and prefers to exist in the cleaved state. This His-tag free NirG was stable and was purified by gel filtration and used in subsequent investigations. Therefore, NirG was the only one of the Nir proteins purified to homogeneity allowing it to be investigated for crystal formation in structural studies.

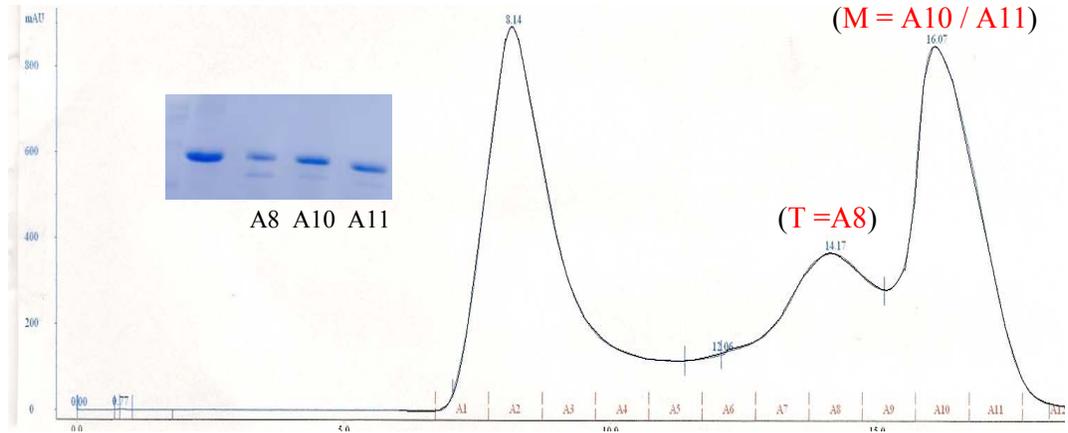
#### Figure 4.11: SDS-PAGE of NirG in HEPES

The first lane is the marker. The other lanes represent each 2 ml fraction of NirG purified by metal chelate chromatography, from a 1L culture. NirG is more stable in 50 mM Hepes, 0.1M NaCl, allowing a sufficient level of NirG to be obtained for further analysis.



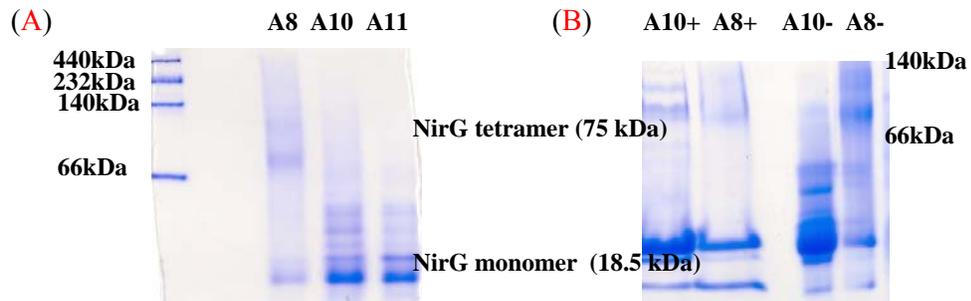
**Figure 4.12: Elution profile of NirG in HEPES**

NirG purified in HEPES has a markedly different elution profile in comparison to the elution profile for NirG in 50 mM Tris pH8, 0.5M NaCl (Figure 4.5). Here we see two forms of NirG being eluted, as a tetramer (T) and a monomer (M). These fractions were subsequently analysed on non-denaturing (native) gels. The inset is an SDS-PAGE gel of NirG from a His-bind column and fractions A8, A10 and A11 from gel filtration.



**Figure 4.13: Non-denaturing (native) gels of NirG of the corresponding fractions from gel filtration chromatography**

A8 and A10/A11 correspond to the fractions eluted from the gel filtration column as per the elution profile above for the tetramer and monomer respectively. (A) a native gel of fractions A8, A10 and A11. (B) a native gel run with A8, A10 and A11 in the presence (+) and absence (-) of  $\beta$ -mercaptoethanol as indicated, whose presence reveals that the tetramer is held by disulphide bonds.



#### 4.3.4. NirG crystallisation trials

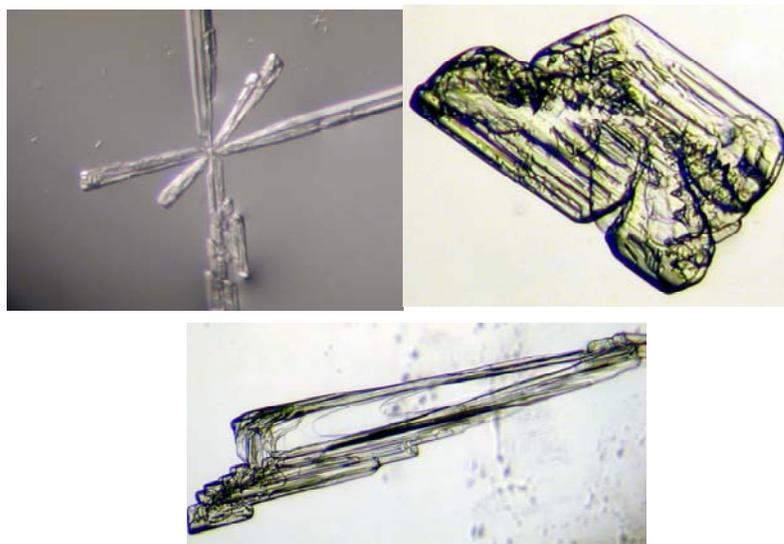
Of NirD, NirL, NirG and NirH, only NirG was purified to homogeneity and sufficient stability for investigation of crystal formation to deduce the structure from X-ray crystallography. This would enable us to obtain a precise description of the tertiary structure of NirG and allow comparison of NirG to proteins of known structure and similarity. This would make possible the assignment of a role and function for NirG.

The hanging drop method vapour diffusion method was used to screen for optimal conditions for crystallisation. Optimal conditions for NirG crystals was obtained at 60 % 2-methyl-2, 4-pentanediol (MPD), 0.1M Tris pH 8.5 and 0.2M ammonium phosphate after much analysis of various precipitant solutions. There was no dependency on temperature, as the screens were performed at 4°C as well as 16°C. The various crystal formations are shown in Figure 4.14.

Unfortunately, analysis by Dr Brindley revealed these to be salt crystals and not protein crystals.

**Figure 4.14: Crystal formations in 60 % MPD, 0.1M Tris pH 8.5 and 0.2M ammonium phosphate.**

Pictures of the crystal formations obtained by screening for NirG with various conditions. These were later identified to be salt.



#### 4.4. NirF

Conflicting arguments exist regarding the precise role of NirF, a protein of 43.3 kDa. Whilst Palmedo *et al* suggest the similarity of NirS and NirF (41 % nucleotide identity) indicate a scaffold role for a precursor of haem  $d_1$  or dehydrogenase activity with PC2 to yield SHC (Palmedo, Seither *et al.* 1995), other papers suggest a role in ferrochelation (Kawasaki, Arai *et al.* 1995). Indeed, a similarity to a cobalamin decarboxylase, CobT, also exists (Kawasaki, Arai *et al.* 1995; Zumft 1997). It is believed to be translocated to the periplasm due to the presence of a signal in *Pa. denitrificans* NirF, but this is not clearly observed with NirF from *Ps. aeruginosa* and *Ps. stutzeri*. However, a high level of sequence similarity at an average of 55 % of NirF exists between the species suggesting they share a periplasmic location (Table 4.2) (Baker, Saunders *et al.* 1997). This diminishes the theory that NirF acts as a scaffold, as from our previous chapter, NirE is clearly a cytoplasmic protein and therefore if it were to act as a scaffold for the assembly of haem  $d_1$  it would not reside in the periplasm.

Bioinformatic analysis shows little similarity to any proteins of known structure and function. The similarity of NirF to the  $d_1$  domain of cytochrome  $cd_1$  has implied a  $\beta$ -propeller domain structure (Baker, Saunders *et al.* 1997), and a role for binding haem  $d_1$ . This also suggests that NirF may have evolved as a result of gene duplication, but it is unknown whether its necessity in haem  $d_1$  biosynthesis is by binding it in a  $\beta$ -propeller domain, or through the synthesis of intermediates. This remains to be elucidated through discovery of its structure and function attempted in this thesis.

Furthermore, NirF has 21 % identity to NirE in the first 100 N-terminal residues, suggesting a similar methyltransferase role but the SAM binding site is somewhat different in NirF to NirE, suggesting a possible interaction with SAM (Figure 4.15). However, in Chapter 3, experiments with a *S. enterica cysG metH* mutant deficient only in its methyltransferase activity, with a plasmid containing NirF was not able to complement the cysteine auxotrophy of this mutant. Therefore we can say with confidence that NirF confers no such activity and NirE is the only SUMT of the *nir* operon.

NirF was accordingly overproduced in *E. coli*, in the same way as NirD, NirL, NirG, and NirH in the previous section. An increase in the level of protein is

observed with induction of protein overproduction at 16 °C overnight (17.2 mg/L) in comparison to 37 °C (9.3 mg/L). Though NirF was found to be soluble, it is highly unstable and rapidly degraded as observed by SDS-PAGE analysis (Figure 4.16A). The addition of protease inhibitors markedly changes the stability of NirF (Figure 4.16B), but analysis of NirF by gel filtration chromatography shows it is mostly aggregated (Figure 4.17). Homologous expression of NirF in *Ps. aeruginosa* revealed the same results as observed with NirD, NirL, NirG, and NirH obtained aerobically and anaerobically. NirF is not well over produced aerobically and did not increase with the induction of anaerobiosis in *Ps. aeruginosa* (Figure 4.18). Decreasing the concentration of IPTG in Tuner cells to 0.1 mM and 0.04 mM did not alter the yield of soluble NirF (data not shown). NirF was also screened for stability in a variety of buffers as detailed in Section 4.2.5, but was to no avail, as was the addition of DTT (data not shown). Therefore, work is still required to overproduce NirF with a sufficient level of stability to allow for further characterisation.

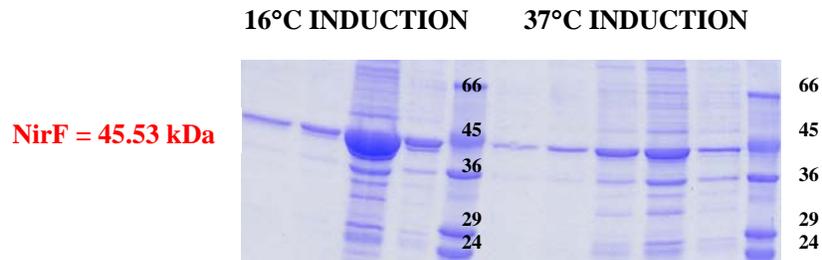
#### Figure 4.15: Comparison of sequences of NirF and NirE

An alignment of NirF and NirE from *Ps. aeruginosa* shows a 21 % identity in the first 100 N-terminal residues. The asterisks show identical amino acids, dots those belonging to the same group, and the SAM binding site is shown in red.

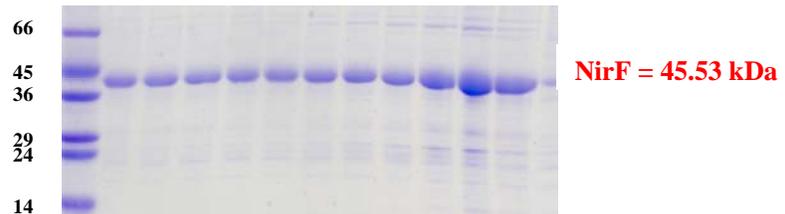
NirE	MNTTVIPPSLLDVFDPAGSVALV	<b>GAGPG</b>	DPGLLTLRAWALLQQA	EVVVYDRLVARELIAL	60
NirF	MNLRPLAPLLLTLLAGCSQQPPL-	RGSGDLGVLI	ERADGSVQILDGTAKTSLARVEGLGD	59	
	**	:.*	**	:	...
				:	*.*** **
				:	** . :*
				:	.. *. * :.
NirE	LPESCQRIYVGKRCGHHSLPQEEINELLVRLARQRRVVR	100			
NirF	LSHASLVFSRDQRYAYVFGRDGGLTKDLLAQRIDKRLIQ	119			
	*..:	:	.*	..	:
		:	.*	:	..:*
		:	.*	:	* :*:::

**Figure 4.16: SDS-PAGE analysis of NirF with induction at 16 °C and 37 °C and in the presence of protease inhibitors**

(A) Optimal protein yield was achieved from induction at 16 °C overnight, in comparison to 37 °C overnight. The SDS-PAGE shows the fractions eluted at 400 mM imidazole. Cultures and samples were treated identically except for the different incubation temperature.

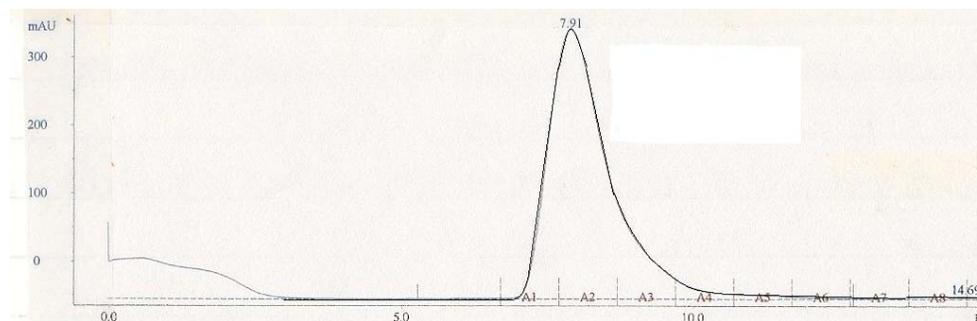


(B) NirF is more stable and less prone to degradation in the presence of protease inhibitors, as shown by SDS-PAGE of eluted fractions from a His-bind column.



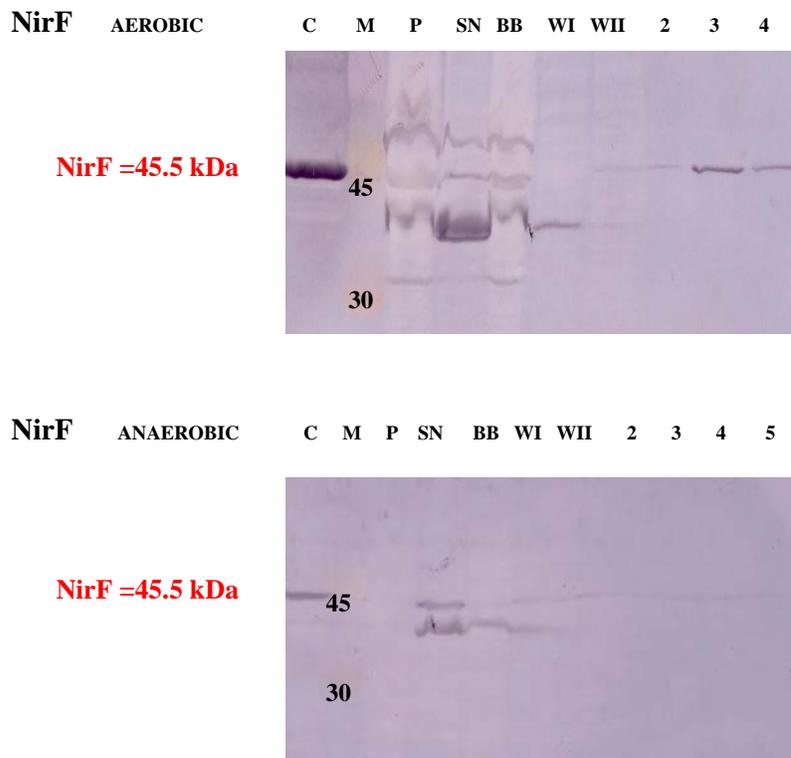
**Figure 4.17: Elution profile of NirF with gel filtration chromatography**

The resultant elution profile of NirF shows it is eluted as aggregated protein.



**Figure 4.18: NirF western blots from overproduction in *Ps. aeruginosa* aerobically and anaerobically**

NirF overproduction is not clearly visible on SDS-PAGE, therefore western blot analysis was performed by raising antibodies to the His-tag. **C** is the control purified from *E.coli*. **M** is the protein marker. **SN** represents the supernatant and **P** the pellet after sonication. **BB**, **WI**, **WII** are washes with buffer of increasing imidazole (5 mM, 50 mM, 100 mM respectively). The numbered lanes represent purified protein eluted at 400 mM imidazole. NirF has a lower level of overproduction in comparison to the control, from *E. coli*. No difference in protein level is observed with aerobic or anaerobic growth.



#### 4.5. NirJ

Bioinformatic analysis of the *Ps. aeruginosa* NirJ sequence (Figure 4.19) shows the presence of a conserved N-terminal motif for radical SAM enzymes. This motif, CXXXCXXC, predicts it to be a [4Fe-4S] protein, similar to that observed in MoaA, NifB, HemN and PqqE proteins for which a metal redox-binding site (Fe-S) is postulated. These are all believed to take part in the conversion of SAM to a radical form, being classified as part of the radical SAM protein family (Sofia, Chen et al. 2001). However, the functions of MoaA, NifB, and PqqE are largely unknown and cannot be used to ascertain the precise function of NirJ. It is notable that radical SAM enzymes are involved in anaerobic or oxygen independent mechanisms, which are otherwise performed aerobically with other proteins. Therefore, NirJ is likely to be involved in radical chemistry such that it may be the likely candidate to catalyse the introduction of the oxo groups on the haem  $d_1$  macrocycle, under anaerobiosis. HemN, a coproporphyrinogen III oxidase, is one such enzyme catalysing the anaerobic oxidative decarboxylation of coproporphyrinogen III to yield protoporphyrinogen IX in the haem biosynthetic pathway (Troup, Hungerer et al. 1995; Layer, Kervio et al. 2005). HemN converts the propionate side chains at C3 and C8 to the corresponding vinyl groups via a reduction of the Fe-S cluster, electron transfer to SAM, SAM cleavage, and formation of the deoxyadenosyl radical. This radical is then presumed to abstract a hydrogen atom from the substrate propionate side chain, yielding 59-deoxyadenosine and a substrate radical. CO<sub>2</sub> release and vinyl group formation produces protoporphyrinogen IX (Layer, Kervio et al. 2005). For many years the presence of an adenosyl radical has been presumed but has only been confirmed recently by work performed on *Pa. pantotrophus* NirJ (data unpublished, communication Dr Brindley). NirJ from *Pa. pantotrophus* is available at high concentrations and stable for further analysis. It was seen to contain iron sulphur centers and exhibited the characteristic brown appearance that his moiety endows the cells with. EPR investigations of NirJ resulted in the appearance of a radical signal, which is due to the presence of a deoxyadenosyl radical, confirming the mechanism of action of radical SAM enzymes. In the same vein, we may make the assumption that NirJ from *Ps. aeruginosa* will act similarly as it shares 47 % identity with *Pa. pantotrophus* NirJ. Therefore, it may also be inferred that NirJ, in a similar fashion to HemN, is involved in radical chemistry whereby the addition of the oxo groups at C3 and C8 is facilitated. This however remains to be confirmed experimentally.

**Figure 4.19: Multiple sequence alignment of NirJ**

Multiple alignment of NirJ from *Ps. aeruginosa* (PAO) with NirJ *Pa. pantotrophus* (PAPANT), HemN from *Ps. aeruginosa*, HemN from *E. coli* and MoaA from *Ps. aeruginosa*. Asterisks show identical amino acids, dots belonging to the same group. The conserved radical SAM enzyme motif, CXXXCXXC, is shown in red.

```

NirJ_PAO      ----MLRISHYLRSLTEDAPTPRRAG-----GRRAPVVIWNL 33
NirJ_PAPANT  ----MHQLLDPSPPRRRSRP-----DAVRPVVIWNL 27
HemN_PAO     -MLDTIRWDADLIRRYDLSGPRYTSYPTAVQFHEGIGPFDQLHALRDSRKAGHPLSLYVH 59
HemN_E.coli  MSVQQIDWDLALIQKYNYSGPRYTSYPTALEFSEDFGEQAFLOAV--ARYPERPLSLYVH 58
MoaA_PAO     -----MTDWIDGQG-----RQIDYLRLSV 19
              :      :

NirJ_PAO     LRRCNLTCKHCYS-----TSADSDFRGELETAEILRGIDDLRAAGVRVLILSGGEP-- 84
NirJ_PAPANT  TRSCNLKCRHCYT-----VSADRPFPGELSHDQAMAVLRDLSDFRIPALILSGGEP-- 78
HemN_PAO     IPFCANICYCACNKVITKDRGRSAPYLARLVREIEIVSRHLSREQVVEQLHFGGGTPTF 119
HemN_E.coli  IPFCHKLCYFCGCNKIVTRQQHKADQYLDALQEIVHRAPLFAG-RHVSQHLHWGGGTPTY 117
MoaA_PAO     TDRCDFRCVYCYMA-----EDMRFLLPRQVLTLEEIERVARLFVAGVVRKRLRLTGGE 71
              *  *  * * * * * * * * * *

NirJ_PAO     LMHPDLFEIAAHARQ-----AGM-FVALSSNGTLIDEGNIQRVAEARFDYVIGISLDGLR 137
NirJ_PAPANT  MSRFDWFELAAEARR-----LDFRHLSTNGTKIDAGNVERLAGLGFYVIGISLDGIG 132
HemN_PAO     LSPGQLRELMSQLRTHLNLDDSDGYGIEIDPREADWSTMGLLRELGFNRVSLGVQDFD 179
HemN_E.coli  LNKAQISRLMKLLRENQFNAD--AEISIEVDPREIELDVLHLRAEAGFNRLSMGVQDFN 175
MoaA_PAO     LVRPGIVGLCERLAALP-----GLRELCMTSNGSQLVR-YARPLYDAGLSRLNLSLDTLD 125
              :      :      :      :      :      :      :      :      :

NirJ_PAO     ETHDRFRQKQGSFDAALAAMRLCREADIR--VGMRTTTEENAAQLPALDLMRELDVQ 194
NirJ_PAPANT  AVNDWFRGVGEGAFDQALAGVRACKAQGVK--VGLRFTITEGNAHHLFAMLDLCRDEGVD 189
HemN_PAO     MEVQKAVNRMQTPPEETRTIVEAARTLQYRSINLDLIYGLPKQTPDSFARTVDEVIALQPD 239
HemN_E.coli  KEVQRLVNREQDEEFIFALLNHAREIGFTSTNIDLIYGLPKQTPESFAFTLKRVAELNPD 235
MoaA_PAO     PLRFRAITRNGELDKVLAGIDAAQAAGFRRIKLNNAVV-MKGRNADEVPAIVFAIARGLD 184
              :      :      :      :      :      :      :      :      :

NirJ_PAO     KFYLSHLNYSGR--GRRSRALDAHRRRTREALALLFERADQDIRQGR--DSDLFVTGNND 249
NirJ_PAPANT  KFYLSHLVYAGR--GDKHRGEDTEHARTTRAMDLLIARAWQAVERGE--PLEVVTGNND 244
HemN_PAO     RLSVFNYAHLPERFMPQRRINADDLPSPGQKLEMLQRTTEQLAAAGYRYIGMDHFALPDD 299
HemN_E.coli  RLSVFNYAHLPTIFAAQRKIKDADLPSPGQKLDILQETIAFLTQSGYQFIGMDHFARPDD 295
MoaA_PAO     ISFIEEMP-----LGQVGRERGESFCSSEDEVRALIAQRHALLDSAEQ-----SGGPAR 232
              :      :      :      :      :      :      :      :

NirJ_PAO     ADAILLLDWLKRR-RPQQLARLCELLLDWGGNASGEGIANIDNTGEVHPDSYWWHHSVGN 308
NirJ_PAPANT  ADAVYFLRWAETRFAPA VAHLRAHLQAWGGNSSGLGVGNIDPQGRVHPDITYWSDYTLGS 304
HemN_PAO     ELASAQEDGTLQRNFQGYTTHGHCDLVGLGVSAIS-QIGDLYSQNSSDINDYQTSLDNGQ 358
HemN_E.coli  ELAVAQREGVLRNRFQGYTTQGD TDLLGMGVSAIS-MIGDCYAQNQKELKQYYQQVDEQG 354
MoaA_PAO     YVRLVEHPQTRIGFISPHSHNFCATCNRLRLTVEGRLLLCLGHEHSLDLRALRRHPLHD 292
              .      .      :      .      .

NirJ_PAO     IRHQR-----FADFWERPDPLLLQLRQRPRPVVGRCSQCRWLDICNGNTRTRAWAG-G 361
NirJ_PAPANT  VKERP-----FSAIWTG-DDPILATLRTPRPLKGRGCACAYQAVCGGNTRIRALQLTG 357
HemN_PAO     LAIRRGLHCN-SDRVRAVIQQLICHFELAFEDIETEFIDFRSYFAELWPDLERFAADG 417
HemN_E.coli  NALWRGIALTRDDCIRRVDVIKSLICNFRLDYAPIEKQWDLHFADYFAEDLKLAPLAKDG 414
MoaA_PAO     G-----PLLEAIGEALQRKPARHEFSAAAGEVQVLRFMNMSGG----- 329
              .      .      :      .      .

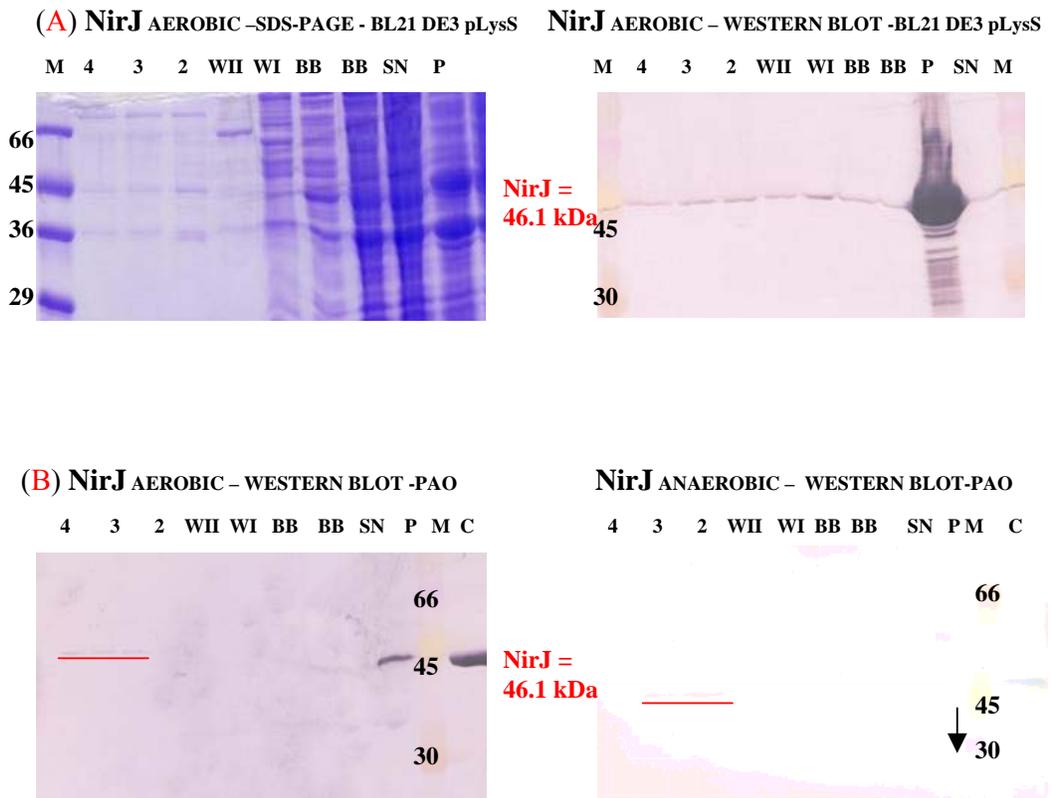
NirJ_PAO     ELWGEDPGCYLSDQ---EIG-----LERIALHAV----- 387
NirJ_PAPANT  DPWAEDPACYLSGS---EIGAEGADLRLAVTPFRGKSHDPAHRFL 400
HemN_PAO     LIRLDAKGIDITSSGRLLVRSICMLFDRYLPSLNR-QRFSRVI--- 459
HemN_E.coli  LVDVDEKGIQVTAKGRLLIRNICMCFDITYLRQKARMQQFSRV--- 456
MoaA_PAO     -----
    
```

During the process of cloning *nirJ*, a database sequence error, R211R, was discovered in all sequencing projects when confirming the cloned *nirJ* was mutation free. This was observed as a change from CGC (Arg) to a database error of CGG (Arg).

Overproduction of NirJ in *E. coli* shows it to be highly insoluble and no change in protein level is observed for NirJ with a change of conditions during induction to 16 °C, instead of 37 °C, as well as varying the length of induction from 3 hours to overnight. The majority of NirJ remains in the cell debris fraction rather than the supernatant of the cell. The presence of a small amount of NirJ is observed in the eluted fractions off the His-bind column only by western blotting in both *E. coli* and *Ps. aeruginosa* (Figure 4.20). Controlling the level of expression by decreasing the concentration of IPTG in Tuner cells had no effect on the yield of soluble NirJ (data not shown). Again, NirJ was screened for solubility and stability in a number of buffers as detailed in Section 4.2.5, but no significant change was observed. NirJ was the least soluble of all the Nir proteins. In comparison to NirJ from *Ps. aeruginosa*, NirJ from *Pa. pantotrophus* is much more stable and soluble. Therefore, further examination is required to obtain NirJ from *Ps. aeruginosa* that is viable for further experimentation.

**Figure 4.20: SDS-PAGE and western blot of NirJ in *E. coli*, BL21 DE3 pLysS and *Ps. aeruginosa* (aerobic and anaerobic)**

(A) SDS-PAGE and western blot analysis of NirJ shows the majority of NirJ is found in the cell debris fraction. A small amount of NirJ is seen in the eluted fractions. **M** is the protein marker. **SN** represents the supernatant and **P** the pellet after sonication. **BB**, **WI**, **WII** are washes with buffer of increasing imidazole (5 mM, 50 mM, 100 mM respectively). The numbered lanes represent purified protein eluted at 400 mM imidazole. NirJ is highly insoluble, the majority being drawn into the pellet. (B) No difference in protein level is observed with aerobic or anaerobic growth. NirJ is seen at 46.1 kDa.



#### 4.6. Conclusion

All the potential haem  $d_1$  biosynthetic genes *nirF*, *nirD*, *nirL*, *nirG*, *nirH*, *nirJ*, from *Ps. aeruginosa*, were successfully cloned into pET14b. Preliminary attempts at overproducing the respective proteins in *E.coli* revealed them to be predominantly insoluble remaining largely in the cell pellet after cell lysis and centrifugation. Each was found to have varying degrees of solubility and stability. This was surprising it had been predicted that the Nir proteins should be hydrophilic and thus soluble, without membrane spanning domains (Palmedo, Seither et al. 1995). The instability of the Nir proteins also warranted the inclusion of protease inhibitors in subsequent investigations. However, NirJ was found to be predominantly insoluble being drawn completely in the cell debris fraction. A database error in the nucleotide sequence was also discovered (R211R) during sequencing of NirJ and alignment with the database sequence. Notably, NirD (19.21 kDa) is co purified with another form of NirD (22 kDa) visible on SDS-PAGE, such that they form a very tight complex that cannot be separated by gel filtration. This is the first time that such results have been observed for NirD and may be suggestive of post translation modification and warrants further investigation.

Co-expression of *nirDLGH*, where only NirD is His-tagged revealed that NirD is co-eluted with NirG. This was as predicted from the observed higher complementarity of NirD to NirG as oppose to NirH and NirL. This is the first time that NirD has been experimentally shown to have an affinity towards binding NirG. It has been well hypothesised that the complementarity of NirD, NirL, NirG and NirH is likely to form a multimeric and thus, a multifunctional enzyme (Palmedo, Seither et al. 1995). Therefore it may further be hypothesised that NirL will have an affinity towards binding NirH on the basis of the complementarity that they share. Further examination is required to see if they do indeed form a complex of NirDLGH together.

In an attempt to optimise overproduction of the Nir proteins, the level of induction of IPTG was changed ten fold (0.04 mM) and four fold (0.1 mM) by growth in a cell line that allow homogeneous induction of cell in a culture (tuner cells). The results showed no change in the amount of protein that fell into inclusion bodies and controlling the level of expression by decreasing the concentration of IPTG had no effect on the level of soluble protein for NirF, NirD, NirL, NirG, NirH and NirJ.

A new system for protein expression in *Ps. aeruginosa* was formulated to overcome the problems with solubility and stability of the recombinant proteins when overproduced in *E. coli*. The level of overproduction of recombinant NirF, NirD, NirL, NirG, NirH and NirJ in *Ps. aeruginosa* was observed to be very low under both aerobic and anaerobic conditions and required detection by western blotting. No dependency on the presence or absence of oxygen was observed for overproduction in *Ps. aeruginosa*. This was unexpected as it was predicted that anaerobiosis increases the level of *nir* expression in response to the presence of nitrate and the absence of oxygen, the the level of overproduced protein was also expected to be high (Filiatrault, Wagner et al. 2005).

Since protein levels in *Ps. aeruginosa* were insufficient for further analysis, the overproduced recombinant proteins from *E. coli* were screened for stabilisation in a range of buffers. So far, NirG has been suitably stabilised. The gel filtration elution profile of NirG revealed the presence of a monomer and a tetramer. This study shows for the first time that NirG may indeed exist as a tetramer. This is particularly interesting as it implies that NirD, NirL, and NirH may also be bound to this tetramer to produce a larger multifunctional enzyme, or indeed it implies that NirD, NirL, and NirH with NirG is itself a tetramer, as proposed previously and NirG is displaying its preference to exist in this formation. This remains to be confirmed

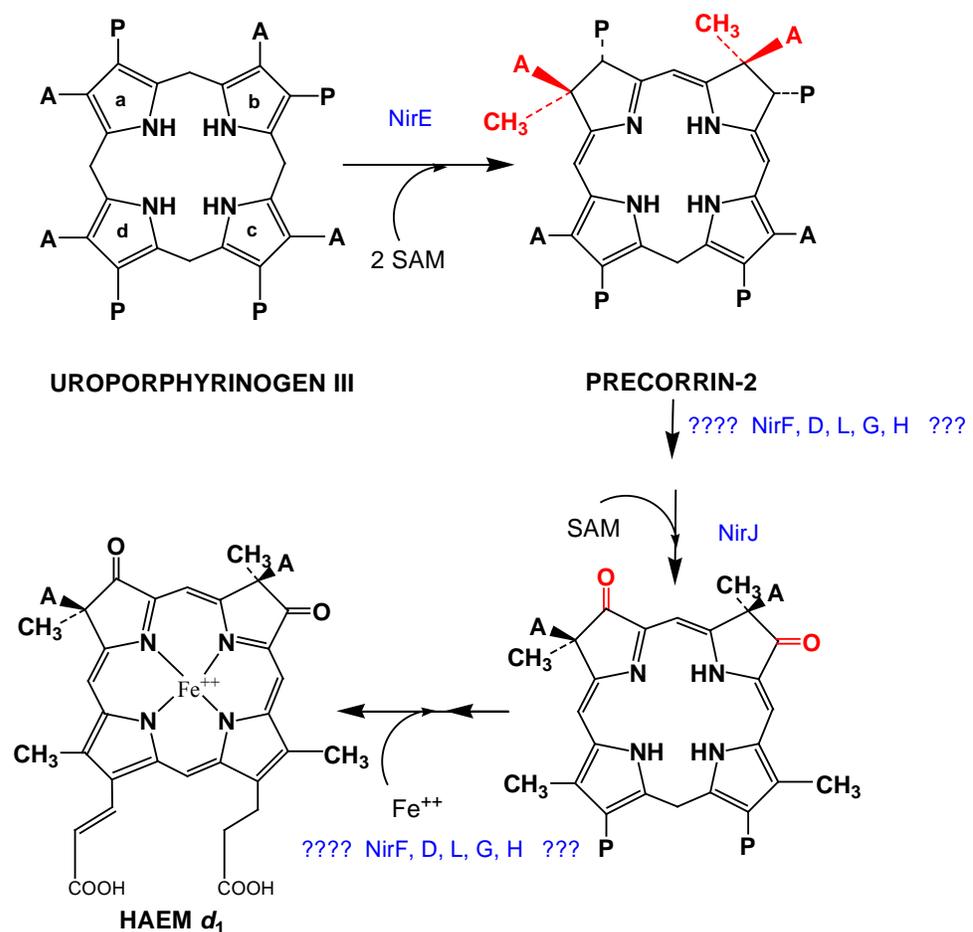
From Chapter 3, NirE catalyses the first step transforming UIII to PC-2. In this chapter, NirJ can be postulated to perform the anaerobic oxidative decarboxylation of the macrocycle by the removal of the propionates and the subsequent addition of the oxo- groups at C3 and C8 in a similar fashion to the reaction catalysed by HemN. The scheme is summarised in Figure 4.21. It is unknown which of the remaining steps NirF, NirD, NirL, NirG and NirH perform to haem  $d_1$ .

Whilst assignments of role or function cannot be made for the all the Nir proteins, except for NirJ and NirE, this thesis further endeavours to optimise the stability of these proteins. This would be invaluable in allowing the proteins to be structurally characterised to give further insight into the function of these proteins and possible reactions they may catalyse on the pathway to haem  $d_1$ . Thus enabling a precise

description of tertiary structures of the Nir proteins and further allow the comparison of the Nir proteins to those that are structurally similar with known function. This would make possible the assignment of roles and functions for the Nir proteins that to this point are still a mystery.

### Figure 4.21: The steps of haem $d_1$ biosynthesis

In Chapter 3, NirE is confirmed to catalyse the first step in the methylation of UIII at C2 and C7 to form PC-2. In this chapter NirJ can be postulated in the removal of the propionates and the subsequent addition of the oxo- groups at C3 and C8. The roles of NirF, NirD, NirL, NirG and NirH are still unclear.



**A Synthetic Biology Approach Towards the  
Generation of Tetrapyrrole Intermediates of the  
Haem  $d_1$  Pathway, *In Vivo* and *In Vitro***

## 5. A synthetic biology approach towards the generation of tetrapyrrole intermediates of the haem $d_1$ pathway, *in vivo* and *in vitro*.

### 5.1. Introduction

Haem  $d_1$  is a unique tetrapyrrole found only in bacterial dissimilatory nitrite reductases, acting as the active centre for nitrite reduction. It differs from other known porphyrins, such as haem  $b$ , being more saturated and characterised by the presence of two oxo groups at C3 and C8, methyl groups at C2 and C7 and an acrylate in the C17 peripheral group.

All the genes necessary for haem  $d_1$  biosynthesis have been identified by mutational and complementation analysis in *Ps. aeruginosa* and *Ps. stutzeri* (Palmedo, Seither et al. 1995; Kawasaki, Arai et al. 1997). This thesis identifies a *Pa. denitrificans* operon containing exactly the same genes, but in a different order to *Ps. aeruginosa* and *Ps. stutzeri*, as detailed in Chapter 4. Interestingly, the genes for *nirD* and *nirL* are fused in *Pa. denitrificans*. Essentially the genes required for the production of haem  $d_1$  constitute *nirF*, *nirD*, *nirL*, *nirG*, *nirH*, *nirJ*, and *nirE* forming an operon.

Thus far, a function has only been confirmed for NirE, as a SUMT catalysing the SAM dependent transformation of UIII to PC-2 as proven in Chapter 3. The rest of the pathway from PC-2 to haem  $d_1$  remains uncharacterised. To date no true intermediates have been isolated and as a result there exists a distinct lack of information as to how the aforementioned atypical substituents are introduced into the macrocycle. The most intriguing of which is the incorporation of oxygen onto the macrocycle under anaerobic conditions. In comparing their structures, shown in Figure 5.1A, the synthesis of haem  $d_1$  from PC-2 requires the decarboxylation of the acetate side chains at C12 and C18; loss of the propionic side chains at C3 and C8 with subsequent oxidation at C3 and C8; dehydrogenation of the C17 propionate side chain to give the acrylate substituent and ferrochelation.

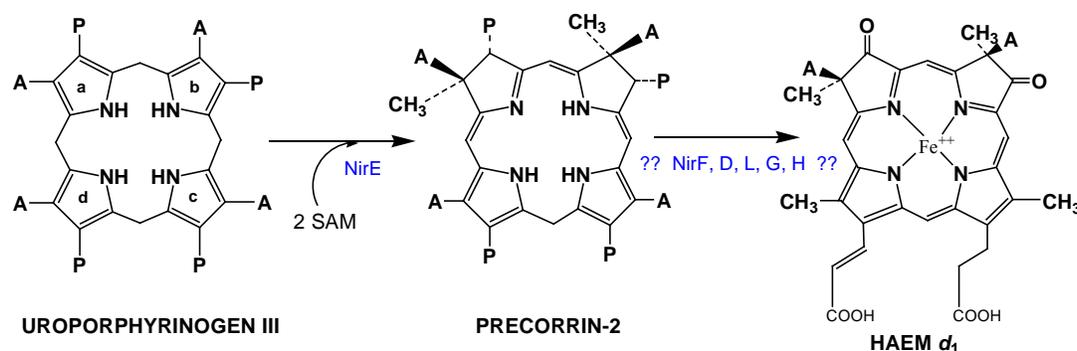
Only one putative intermediate has been put forward, termed Compound 800 (Figure 5.1B), after the unique mass of its methyl ester derivative, where cell extracts of *Ps. aeruginosa* with NirFDLGH and SUMT from *D. vulgaris* were incubated anaerobically with SAM and UIII (Youn, Liang et al. 2004). There is no inclusion of NirE, the preferred SUMT for this pathway, and also NirJ, which is

shown to be necessary for haem  $d_1$  biosynthesis (Palmedo, Seither et al. 1995; Kawasaki, Arai et al. 1997). Treatment of C800 with strong acid and heat results in haem  $d_1$ , however these are non-physiological conditions that will not occur otherwise *in vivo*. C800 has some unusual substituents such as epoxide and sulfoxide moieties that have not previously been found in any tetrapyrrole intermediate. Therefore, it remains to be seen if C800 is a true intermediate or related to an intermediate in the pathway to haem  $d_1$ .

### Figure 5.1: The pathway from UIII to haem $d_1$

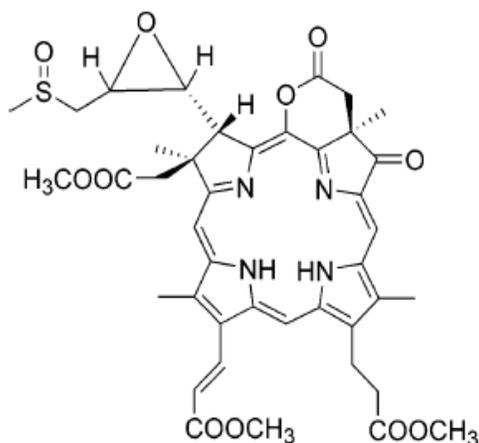
#### A: The synthesis of haem $d_1$ from UIII via PC-2

The synthesis of haem  $d_1$  from PC-2 requires the decarboxylation of the acetate side chains at C12 and C18; loss of the propionic side chains at C3 and C8 with subsequent oxidation at C3 and C8; dehydrogenation of the C17 propionate side chain to give the acrylate substituent and ferrochelation



#### B: The structure of Compound 800.

The structure of C800 bears a resemblance to haem  $d_1$  in the southern half of the macrocycle, whereas the sulfoxide and epoxide moieties are in the northern half.



In light of the lack of any characterised intermediates available, this chapter explores the journey in the engineering of a strain to synthesis haem  $d_1$  *de novo*, and thus any intermediates in the pathway. This is performed in a number of ways. Primarily, bacteria were engineered with plasmids containing the *nir* genes in predefined arrangements and combinations, in the hope that these would produce haem  $d_1$  or any of the preceding intermediates *in vivo*. This was performed in both *E. coli* and *Ps. aeruginosa* in a number of ways. The next method used soluble cell extracts with predefined overproduced Nir proteins either singularly, or combinations thereof, to check for activity towards PC-2 and SHC, synthesised *in vitro*. The presence of key cofactors was also varied. The data obtained are discussed herein.

## RESULTS AND DISCUSSION

### 5.2. *In vivo* accumulation of intermediates

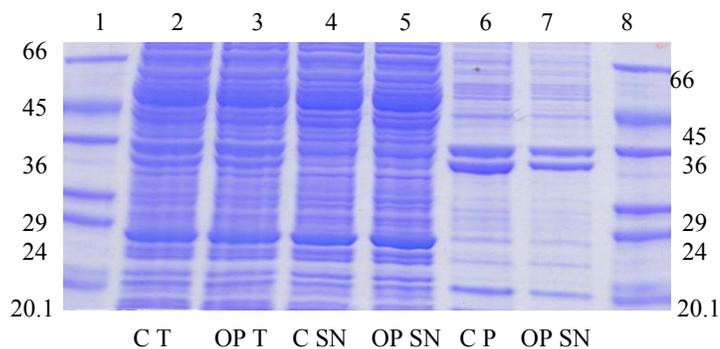
It was highly attractive to engineer a bacterial system with the *nir* operon to synthesise haem  $d_1$ , and in this manner too, synthesise any intermediates *in vivo* by defining the genes or combinations thereof to be overexpressed. This would not only negate the problems of instability and the amount of soluble purified protein, but also show which reaction or reactions they may be catalysing *in vivo* in both *E. coli* and *Ps. aeruginosa*. This method of tetrapyrrole intermediate production has been performed with success before and given rise to previously uncharacterised intermediates in the vitamin B<sub>12</sub> pathway (McGoldrick, Roessner et al. 2005).

#### 5.2.1. The expression of the operon in *E. coli*

A plasmid containing the haem  $d_1$  operon cloned into pKK233.3 with *nirFDLGHJE* all in succession was kindly provided by Dr Raux-Deery. Preliminary attempts to overexpress the genes were unsuccessful. This was performed in a number of strains by transforming the plasmid into JM109, ER171 (CysG deleted strain) and CysG<sup>-247</sup> (CysG deleted strain, but with CobA). These were then cultured in 1L volumes and grown with IPTG induction. Controls containing no plasmid were also grown at the same time and subject to the same conditions. The resultant cell extract and pellets obtained after sonication and centrifugation were then analysed for Nir protein over production by SDS PAGE, shown in Figure 5.2. The cell extract and the pellets exhibited no fluorescence indicative of no tetrapyrrole compounds except for those expressed in CysG<sup>-247</sup>, due to the presence of CobA. Analysis of the SDS-PAGE confirms this since there is no difference between the control and the strain expressing the plasmid. No proteins of predicted size for NirF, D, L, G, H, J or E is observed. This result is likely to be due to poor translation of the protein as no such factors were included in the cloning between the genes and they were cloned straight after each other. Therefore, the expression of the operon must include this.

**Figure 5.2: SDS-PAGE of the haem  $d_1$  operon overexpressed in *E. coli*.**

An SDS-PAGE of the haem  $d_1$  operon expressed in JM109 alongside a control of JM109 only is shown. Clearly no difference can be seen between our control and the strain overexpressing the operon. Lane 1 is the marker, lane 2 control total cell fraction (CT), lane 3 the total cell fraction with the operon (OP T), lane 4 the control isolated supernatant (C supernatant), lane 5 isolated supernatant with the operon (OP SN), lane 6 the control cell pellet (C P) and lane 7 the cell pellet with the operon (OP SN). The results for overexpression in ER171 and CysG<sup>-</sup>247, showed no difference from the controls either.

**5.2.2. Link and Lock cloning of *nir* genes into an operon**

The problems in expressing the operon were apparent, as we neither saw the accumulation of tetrapyrroles nor the overproduction of the Nir proteins. To overcome this, the operon was cloned with the incorporation of ribosome-binding sites (RBS) before each gene to allow for greater translation.

The cloning of the *nir* operon was achieved by the Link and Lock method of cloning, whereby genes may be consecutively cloned in to a plasmid, in order, one after the other by utilising the ability of neoisoschizomer restriction enzymes to form compatible cohesive ends. These when ligated, do not generate the same restriction sites and therefore can be re-used in introducing other genes, under the proviso that they do not occur naturally within the genes or plasmid.

In this study *NheI* (GCTAGC) and *XbaI* (TCTAGA) were used, both forming compatible cohesive ends that when ligated are no longer restrictable (Section 2.2.12). These sites are not found within the *nir* genes or the plasmid, pUCPNde. Primers were designed for each of the genes *nirE*, *nirJ*, *nirG*, *nirH*, *nirL*, *nirD* and *nirF*, to introduce, *NdeI* and a RBS (AGGAG(A)) for the first gene, *nirE* or *nirF*,

and then *NheI* and RBS at the 5' end for all other genes. *XbaI* and *BamHI* restriction sites were introduced at the 3' termini of all the *nir* genes. The introduction of these sites allowed Link and Lock cloning of the operon into the pUCPNde expression vector allowing the encoded proteins to be overproduced both within *E. coli* (JM109) and wild type *Ps. aeruginosa* (PAO1).

The first gene in the sequence, for example *nirE*, was cloned into the plasmid using *NdeI* at the 5' end and *XbaI* – *BamHI* at the 3' end. The next gene and all subsequent genes, pre-restricted with *NheI* and *BamHI*, are then cloned consecutively into the plasmid that is prerestricted with *XbaI* and *BamHI*. The compatibility of *XbaI* and *NheI* allows ligation but these are no longer restrictable after ligation, therefore they can be reused to clone genes into the plasmid in subsequent steps. Please refer to Section 2.2.12 for a detailed account. The strategy is outlined in Figure 5.3.

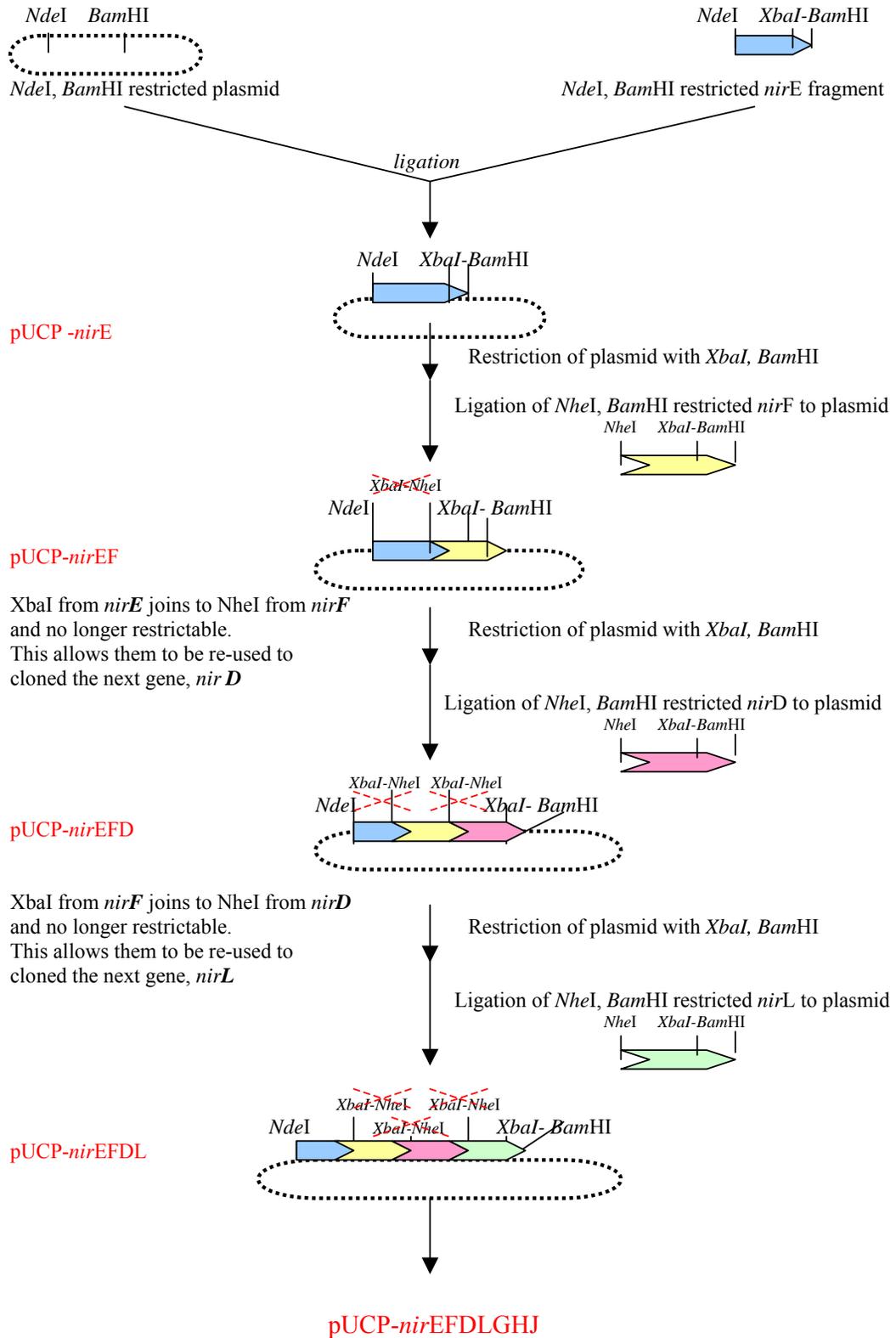
In this manner the entire *nir* operon was cloned into pUCPNde to allow overexpression and the formation of haem  $d_1$  in both *E. coli* and wild type *Ps. aeruginosa*. Accordingly a library of different constructs was created to allow the accumulation of intermediates in both organisms, these are listed in Table 5.1. Two sets of plasmids were created, the first starting with our first gene in the pathway *nirE*, with *nirFDLGHJ* in succession and the second starting with *nirF*, with *nirDLGHJE* in succession. The former allows for an increased level of substrate PC-2 as those genes closest to the promoter tend to be better expressed. The latter to identify if they are required in the same order as in the *Ps. aeruginosa* operon, and thus a dependency on the presence of the other proteins NirFDLGHJ to be expressed first and subsequently a spontaneous reaction occurs as soon as NirE forms PC-2, though there may be less substrate than the former set. The operon is shown to be expressed in Figure 5.4 by SDS-PAGE analysis.

**Table 5.1: Plasmids constructed for the accumulation of haem  $d_1$  intermediates**

SET 1	SET 2
<i>nirE</i> (pVP49)	<i>nirF</i> (pVP58)
<i>nirEF</i> (pVP50)	<i>nirFD</i> (pVP59)
<i>nirEFD</i> (pVP51)	<i>nirFDL</i> (pVP60)
<i>nirEFDL</i> (pVP52)	<i>nirFDLG</i> (pVP61)
<i>nirEFDLG</i> (pVP53)	<i>nirFDLGH</i> (pVP62)
<i>nirEFDLGH</i> (pVP54)	<i>nirFDLGHJ</i> (pVP63)
<i>nirEFDLGHJ</i> (pVP55)	<i>nirFDLGHJE</i> (pVP65)

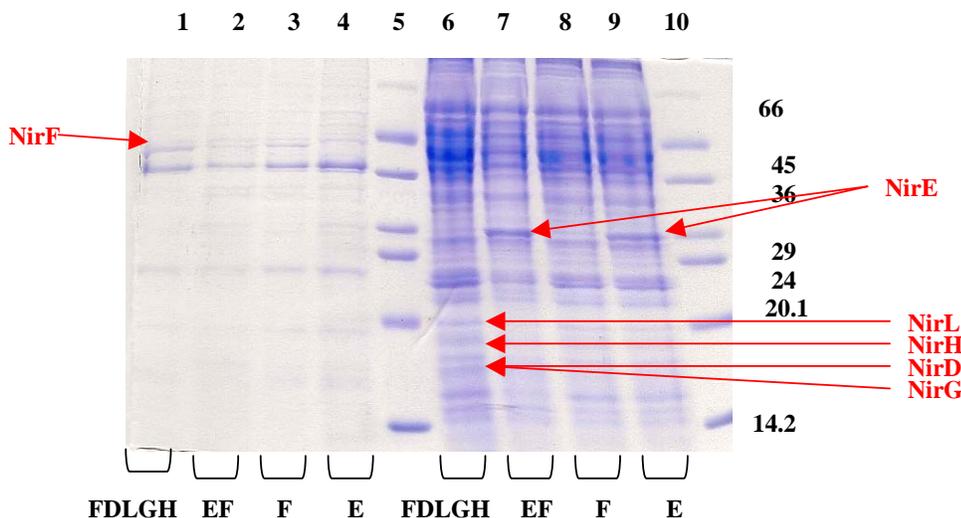
**Figure 5.3: Link and Lock method of cloning**

Diagrammatic representations of the link and lock method of cloning, for a detailed description, please see the text.



**Figure 5.4: SDS – PAGE analysis of the expression of the operon constructed with ribosome binding sites (RBS)**

An SDS-PAGE of the haem  $d_1$  operon expressed in *Ps. aeruginosa*. Lanes 1 to 4 are of the pellet after sonication of NirFDLGH (1), NirEF (2), NirF (3), and NirE (4). Lane 5 is the marker. Lanes 5 to 9 are of the isolated supernatant after sonication, NirFDLGH (6), NirEF (7), NirF (8), and NirE (9). Lane 10 is the marker. The expression of the operon is evident by the presence of all the predicted proteins of correct size for each lane. NirE is 29 kDa, NirF is 44 kDa, NirD is 17 kDa, NirL is 20.2 kDa, NirG is 16.5 kDa and NirH is 19.5 kDa.



**5.2.3. Production, extraction and separation of accumulated tetrapyrroles**

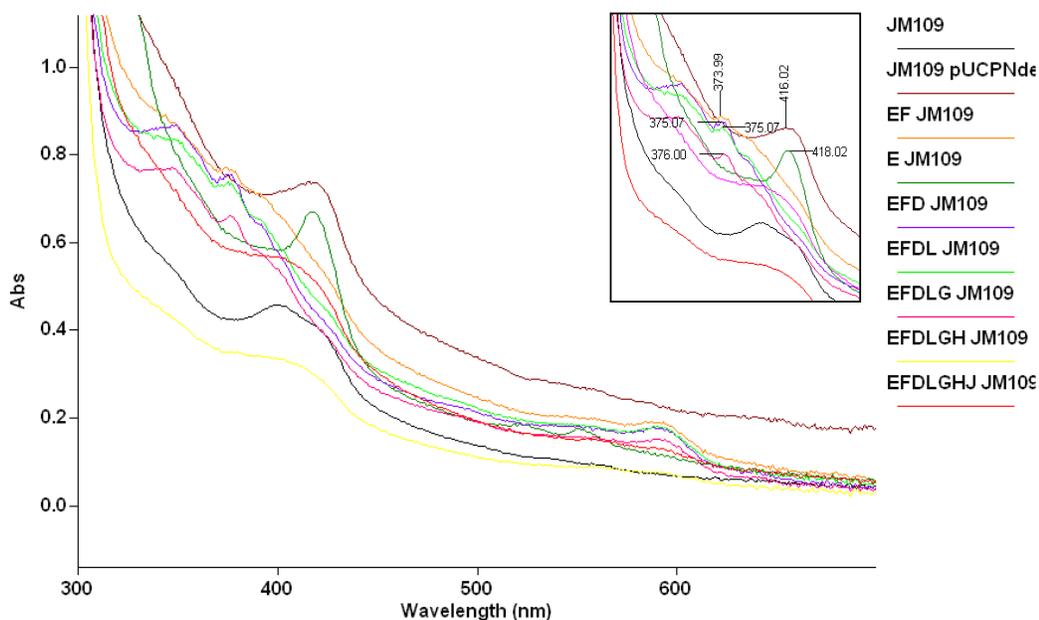
Plasmids containing the *nir* gene operon or combinations thereof shown in Table 5.1, were transformed into *E. coli* JM109 and *Ps. aeruginosa* wild type separately as outlined in Section 2.2.7. 1L cultures were grown either aerobically or anaerobically (with 25 mM sodium nitrate) as appropriate, with induction at 1 mM IPTG. At the same time ALA (10 mg/ml) and SAM (10 mg/ml) was added to the culture. Supplementation of the culture with ALA is known to enhance the production of tetrapyrroles whereby they are secreted into the medium (Harris III 1993). Thus we are able to isolate endogenous tetrapyrroles with ease if they are sufficiently produced. This method is well characterised in both *E. coli* and *Ps. aeruginosa* (Harris III 1993). The media was kept as this stage and the pellets were resuspended in 50 mM Tris-HCl pH 8.0 containing 0.1 M NaCl and sonicated for 30 sec on, 30 sec off three times at an amplitude of 10  $\mu$ n. The media showed no signs of tetrapyrrole accumulation therefore, the resultant soluble cell extract was

applied to RP-18 columns aerobically and anaerobically. The columns were then washed as outlined in Section 2.4.6, and any accumulated compounds were eluted at 25 % ethanol and analysed by a UV-Vis spectrophotometer. The same experiment was repeated, but with separation on DEAE columns, but no difference was observed in the results.

#### 5.2.4. Results obtained from expression in *E. coli* JM109

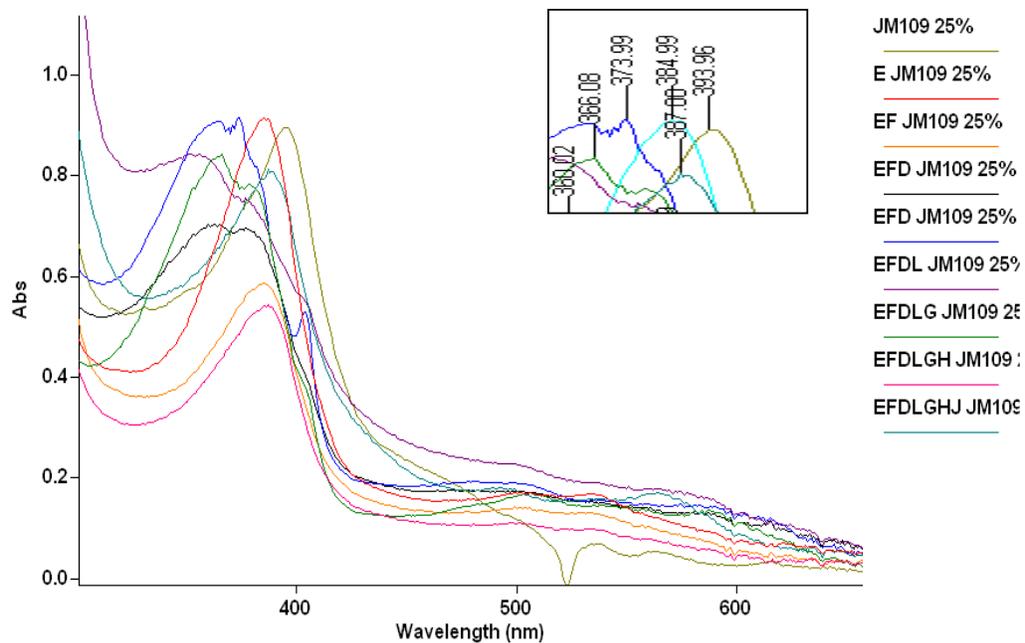
##### Figure 5.5: UV-Vis spectra of all isolated supernatants from *E. coli* prior to separation on RP-18 columns

Overlaid UV-Vis scans of the isolated supernatants of overexpressed *nir* genes in pUCPNde containing *nirE*, *nirEF*, *nirEFD*, *nirEFDL*, *nirEFDLG*, *nirEFDLGH*, *nirEFDLGHJ* and controls JM109, JM109 with pUCPNde only. The inset shows the absorbance maxima for each.



**Figure 5.6: UV-Vis spectra of all isolated fractions from *E. coli* after separation on RP-18 columns**

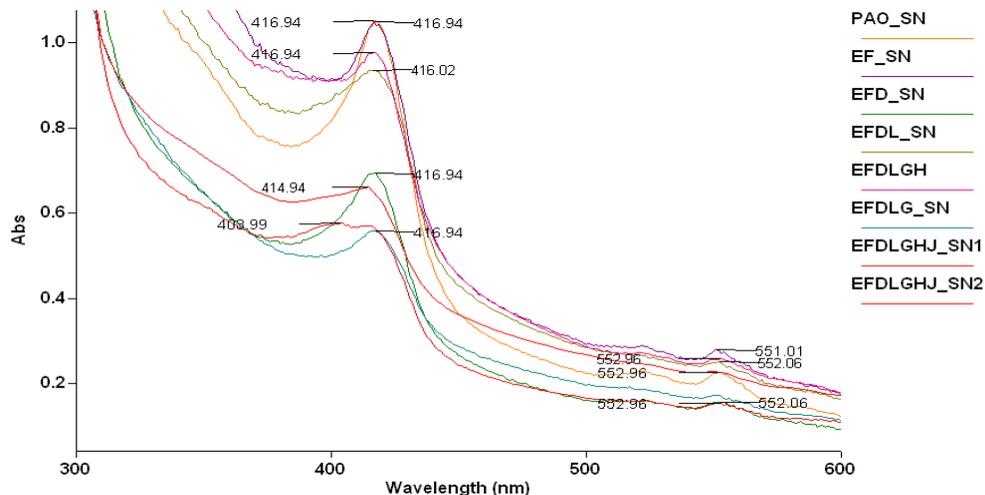
Overlaid UV-Vis scans of isolated fractions from RP18 columns at 25 % ethanol of over expressed *nir* genes in pUCPNde containing *nirE*, *nirEF*, *nirEFD*, *nirEFDL*, *nirEFDLG*, *nirEFDLGH*, *nirEFDLGHJ* and control JM109. The inset shows the absorbance maxima for each.



**5.2.5. Results obtained from expression in *Ps. aeruginosa***

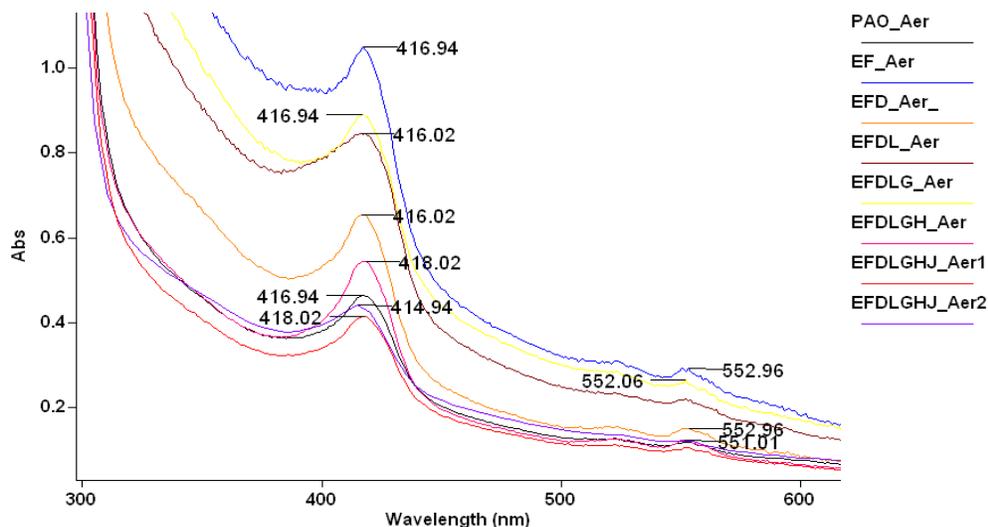
**Figure 5.7: UV-Vis spectra of all aerobically isolated supernatants from *Ps. aeruginosa* prior to separation on RP-18 columns**

Overlaid UV-Vis scans of the isolated supernatants of over expressed *nir* genes in pUCPNde containing *nirEF*, *nirEFD*, *nirEFDL*, *nirEFDLG*, *nirEFDLGH*, *nirEFDLGHJ* and control *Ps. aeruginosa* supernatant. No difference is observed.



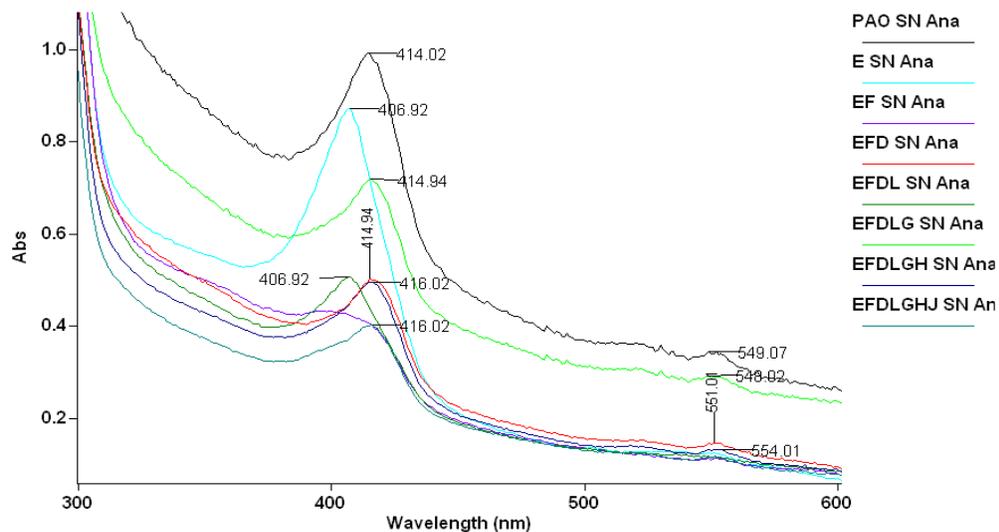
**Figure 5.8: UV-Vis spectra of all aerobically isolated supernatants from *Ps. aeruginosa* after separation on RP-18 columns**

Overlaid UV-Vis scans of the isolated supernatants of over expressed *nir* genes in pUCPNde containing *nirEF*, *nirEFD*, *nirEFDL*, *nirEFDLG*, *nirEFDLGH*, *nirEFDLGHJ* and control *Ps. aeruginosa* supernatant. No difference is observed.



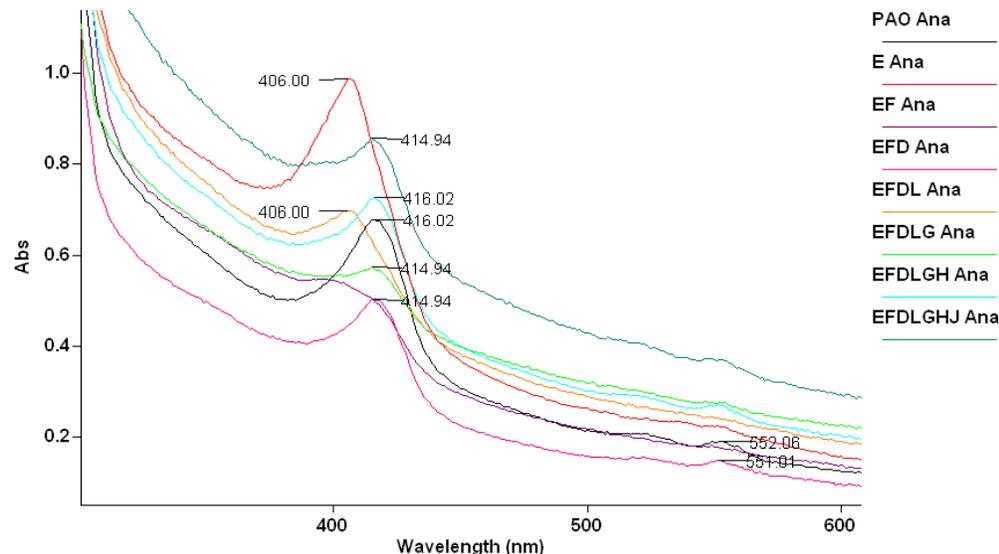
**Figure 5.9: UV-Vis spectra of all anaerobically isolated supernatants from *Ps. aeruginosa* prior to separation on RP-18 columns**

Overlaid UV-Vis scans of the isolated supernatants of over expressed *nir* genes in pUCPNde containing *nirEF*, *nirEFD*, *nirEFDL*, *nirEFDLG*, *nirEFDLGH*, *nirEFDLGHJ* and control *Ps. aeruginosa* supernatant. No difference is observed.



**Figure 5.10: UV-Vis spectra of all anaerobically isolated supernatants from *Ps. aeruginosa* after separation on RP-18 columns**

Overlaid UV-Vis scans of the isolated supernatants of overexpressed *nir* genes in pUCPNde containing *nirEF*, *nirEFD*, *nirEFDL*, *nirEFDLG*, *nirEFDLGH*, *nirEFDLGHJ* and control *Ps. aeruginosa* supernatant. No difference is observed.



### 5.2.6. Discussion of data obtained from expression of the haem $d_1$ operon in *E. coli* and *Ps. aeruginosa*

It was hoped that expression of all the *nir* genes in *E. coli* would endow it with the ability to make haem  $d_1$  since it is not capable of producing it naturally. Similarly, the subsequent homologous expression of the plasmids in *Ps. aeruginosa* was expected to produce haem  $d_1$  more so. Preliminary analysis of the expression of the operon in *E. coli* and *Ps. aeruginosa* showed no evidence of haem  $d_1$  production. The spectrum of haem  $d_1$  is characterised by a solet band at 421 nm, and a band at 608 nm with a weak band at 565 nm and of course has a green hue. No such bands or colours were eluted from the separation of supernatant of *E. coli* and *Ps. aeruginosa* expressing *nirEFDLGHJ* or *nirFDLGHJE* plasmids. No significant differences were observed in this analysis of the expression of *nirF*, *nirFD*, *nirFDL*, *nirFDLG*, *nirFDLGH*, *nirFDLGHJ* and *nirFDLGHJE* (data not shown) and for brevity the rest of the discussion is confined to the expression of plasmids beginning with *nirE*.

The separation of the supernatants of *Ps. aeruginosa* expressing *nirEFDLGHJ* or *nirFDLGHJE* aerobically and anaerobically did not show or bind any substance at all. In fact, it was seen to flow completely through the column as was observed in the analysis of the expression of *nirE*, *nirEF* *nirEFD*, *nirEFDL*, *nirEFDLG*, *nirEFDLGH*, *nirEFDLGHJ* (Figure 5.8 and 5.10), aerobically and anaerobically respectively. This was unexpected, as the genes are being expressed in their native organism, and despite the low level of expression, as seen in Chapter 4, it was assumed that it would be sufficient to produce a small amount of haem  $d_1$  or any intermediates. Furthermore, the expression of the various plasmids under anaerobic conditions in *Ps. aeruginosa* was presumed to be more favourable for haem  $d_1$  or tetrapyrrole production. It is documented that the *nir* genes are up regulated almost 16 fold in response to anaerobic conditions and in particular nitrate, at an optimal concentration of 25 mM (Carlson and Ingraham 1983; Filiatrault, Wagner et al. 2005) Additionally, complete anaerobiosis is not required for expression of the denitrification cluster (Hartig and Zumft 1999). The results obtained may have arisen due to non-exacting conditions and higher requirement for substrates and cofactors that were limiting within *Ps. aeruginosa*. The expression of the *nir* genes was observed to be lower in *Ps. aeruginosa* than in *E. coli* in Chapter 4 and thus may be a limiting factor.

Interestingly, the UV-Vis spectra of each supernatant as shown in Figure 5.5, reveals no significant difference for *nirEFDLGH*, *nirEFDLGHJ* and *nirE* from the controls whereas *nirEF*, *nirEFD*, *nirEFDL* and *nirEFDLG* are showing additional peaks in the region of 375 nm  $\pm$  1.01 nm. This is notable since it may be significant to the presence of SHC whose characteristic spectra shows absorbance maxima at 376 nm and 590 nm. Subsequent separation with RP-18 revealed the presence of a pink band for all the separations except the control but in the case of the *nirEFD*, *nirEFDL* *nirEFDLG*, the band was slightly darker and the spectra showed a broad band with two peaks again in the region of 375 nm but less well defined (Figure 5.6). If these are indeed preliminary data indicating the presence of SHC, it would be advantageous to express the same genes anaerobically and thus perform the separation anaerobically too. It must be noted that the presence of SHC may be a side reaction within the organism and not part of the haem *d*<sub>1</sub> pathway as *E. coli* and *Ps. aeruginosa* possess the capability to produce it *in vivo*. Thus it would also be beneficial to see if this reaction occurs *in vitro* towards PC-2 and a Nir protein or proteins.

Another strategy was explored whereby *Pa. pantrophus nirJ* was to be cloned in the place of *Ps. aeruginosa nirJ* as part of the multigene plasmids. NirJ from *Ps. aeruginosa* is predominantly insoluble in both *E. coli* and *Ps. aeruginosa*. In contrast, NirJ from *Pa. pantrophus* is significantly more soluble and well expressed in *E. coli* (data not published, courtesy of R. Zajicek, Oxford and A. Brindley, Kent). It is notable that *Pa. denitrificans* is unique in comparison as it makes so much NiR that the periplasm preparation is brown and thus, the level of haem *d*<sub>1</sub> synthesis and the enzymes may also be high. However, time was a limiting factor and this could not be completed, but NirJ from *Pa. pantrophus* was used in the production of intermediates *in vitro*, discussed next.

### 5.3. *In vitro* expression of Intermediates

*In vivo* production of haem  $d_1$  and its intermediates has not been successful so far, presumably because of the resultant low steady state concentrations of the Nir proteins within the multigene plasmids. This limitation may be overcome in part by *in vitro* biosynthesis using soluble cell extracts with overproduced proteins and incubating these anaerobically with PC-2 or SHC, alongside probable cofactors.

#### 5.3.1. Production, extraction and separation of accumulated tetrapyrroles

PC-2 and SHC were synthesised *in situ* as described in Section 2.4.2 using isolated soluble cell extracts containing the overproduced proteins supplemented with cofactors for synthesis from ALA (HemB, HemC, HemD, CobA without and with SirC respectively). Alternatively, the same purified proteins were incubated anaerobically with cofactors, in the glove box, generating PC-2 and SHC at higher concentrations, and in a purer form, as described in Section 2.4.3. The spectra for PC-2 and SHC synthesised by this method is shown in Figure 5.16.

Accordingly the *nir* genes were singularly expressed in *E. coli* and the pellet resuspended in 50 mM Tris HCl pH 8.0, 0.1M NaCl. The soluble fractions with the Nir proteins were sonicated and buffer exchanged together in the anaerobic glove box or sonicated separately and buffer exchanged separately and combined in the reaction. This was to allow for any complex formation in case any Nir protein has an affinity to another. Each condition as listed in Table 5.2 was incubated with PC-2 or SHC overnight at 30 °C with or without potential cofactors as appropriate (NAD<sup>+</sup> / NADP / NADPH / SAM / FE II / NaDT). The resultant mixture was applied to a DEAE sephacel column, which had been pre-equilibrated the same buffer in an anaerobic glove box. The columns were then washed as outlined in Section 2.4.7, with increasing concentrations of NaCl, and any accumulated compounds were eluted at 1M NaCl and analysed by a UV-Vis spectrophotometer (HP-352 photodiode array).

**Table 5.2: Summary of incubations of Nir proteins with PC-2 and SHC**

All incubations were performed as described in Section 2.4.4.2 in 50 mM Tris-HCl, pH 8.0, 0.1M NaCl in a total volume of 6 ml. This has been summarised in the table below. Where NirJ was used, it was a mixture of both NirJ from *Pa. pantrophus* as well as *Ps. aeruginosa* NirJ overproduced in *E. coli*. Reaction mixture contained 4.8 ml of cell extract, 1 ml of substrate (PC-2 or SHC), 30 µl each of NAD<sup>+</sup>, NADP, NADPH (10 mg/ml suspension in 50 mM Tris-HCl, pH 8.0, 0.1M NaCl), 20 µl of iron and NaDT (at 1 mM, final concentration is 0.02 mM), made up to 6 ml with 50 mM Tris-HCl, pH 8.0, 0.1M NaCl buffer. **S** = Nir proteins overproduced separately, and sonicated separately before combining in the reaction mixture; **T** = Nir proteins overproduced separately, and sonicated together before adding to the reaction mixture; **All** = All cofactors NAD<sup>+</sup> / NADP / NADPH / SAM / FE II / NaDT, unless otherwise specified.

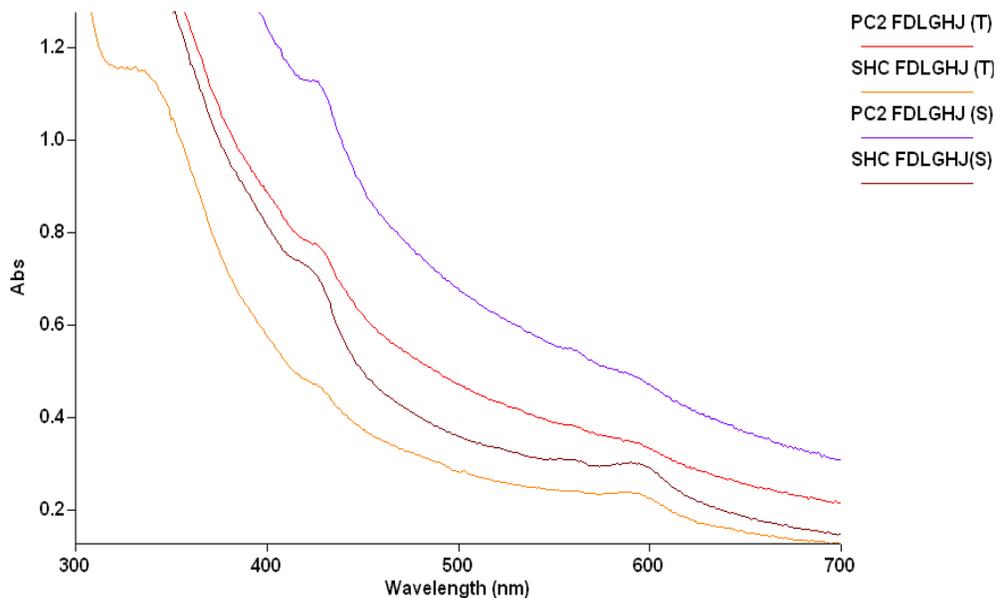
	<b>REACTION</b>	
<b>Incubations with PC-2 and SHC produced <i>in situ</i> from cell extracts</b>		
<b>1.</b>	PC2 + FDLGHJ (T) <i>E. coli</i> + ALL	Figure 5.11
<b>2.</b>	SHC+ FDLGHJ (T) <i>E. coli</i> + ALL	
<b>3.</b>	PC2 + FDLGHJ (S) <i>E. coli</i> + ALL	
<b>4.</b>	SHC + FDLGHJ (S) <i>E. coli</i> + ALL	
<b>5.</b>	PC2 + FDLGHJ (T) + <i>Pa. pantrophus</i> + ALL	Figure 5.12
<b>6.</b>	SHC + FDLGHJ (T) + <i>Pa. pantrophus</i> + ALL	
<b>7.</b>	PC2 + FDLGHJ (S) + <i>Pa. pantrophus</i> + ALL	
<b>8.</b>	SHC + FDLGHJ (S) + <i>Pa. pantrophus</i> + ALL	
<b>9.</b>	PC2 + FDLGHJ (T) <i>Ps. aeruginosa</i> + ALL	
<b>10.</b>	SHC + FDLGHJ (T) <i>Ps. aeruginosa</i> + ALL	
<b>11.</b>	PC2 + FDLGHJ (S) <i>Ps. aeruginosa</i> + ALL	
<b>12.</b>	SHC + FDLGHJ (S) <i>Ps. aeruginosa</i> + ALL	
<b>13.</b>	PC2 + BL21 DE3 pLysS + ALL	Figure 5.13
<b>14.</b>	SHC + BL21 DE3 pLysS + ALL	
<b>15.</b>	PC2 + <i>Ps. aeruginosa</i> + ALL	
<b>16.</b>	SHC + <i>Ps. aeruginosa</i> + ALL	
<b>17.</b>	PC2 + <i>Pa. pantrophus</i> + ALL	
<b>18.</b>	SHC + <i>Pa. pantrophus</i> + ALL	

<b>Incubations of PC-2 with each overproduced Nir protein</b>		
19	PC2 + F	Figure 5.14A
20	PC2 + F + NAD <sup>+</sup>	
21	PC2 + F + ALL	Figure 5.14B
22	PC2 + D + ALL	
23	PC2 + L + ALL	
24	PC2 + G + ALL	
25	PC2 + H + ALL	
26	PC2 + J + ALL	
<b>Incubations of SHC with each overproduced Nir protein</b>		
27	SHC + F + ALL	Figure 5.15
28	SHC + D + ALL	
29	SHC + L + ALL	
30	SHC + G + ALL	
31	SHC + H + ALL	
32	SHC + J + ALL	
33	PC2 + FDLGHJ (T) + ALL (no pD10)	Data not shown
34	SHC + FDLGHJ (T) + ALL (no pD10)	
35	PC2 + FDLGHJ (T) + <i>Pa. pantrophus</i> + ALL (no pD10)	
36	SHC + FDLGHJ (T) + <i>Pa. pantrophus</i> + ALL (no pD10)	
37	PC2 + FDLGHJ (T) + <i>Ps. aeruginosa</i> + ALL (no pD10)	
38	SHC + FDLGHJ (T) + <i>Ps. aeruginosa</i> + ALL (no pD10)	
<b>Incubations with PC-2 and SHC produced from purified protein</b>		
39	PC2 + FDLGHJ + ALL	Figure 5.17A & B
40	PC2 + FDLGHJ (S) + ALL – NO IRON	
41	PC2 + FDLGHJ (S) + ALL – NO NADP/NADPH	
42	SHC + FDLGHJ + ALL	Figure 5.18
43	SHC + FDLGHJ (S) + ALL – NO IRON	
44	SHC + FDLGHJ (S) + ALL – NO NADP/NADPH	

### 5.3.2. Results obtained from incubation of Nir proteins with PC-2 and SHC produced *in situ*.

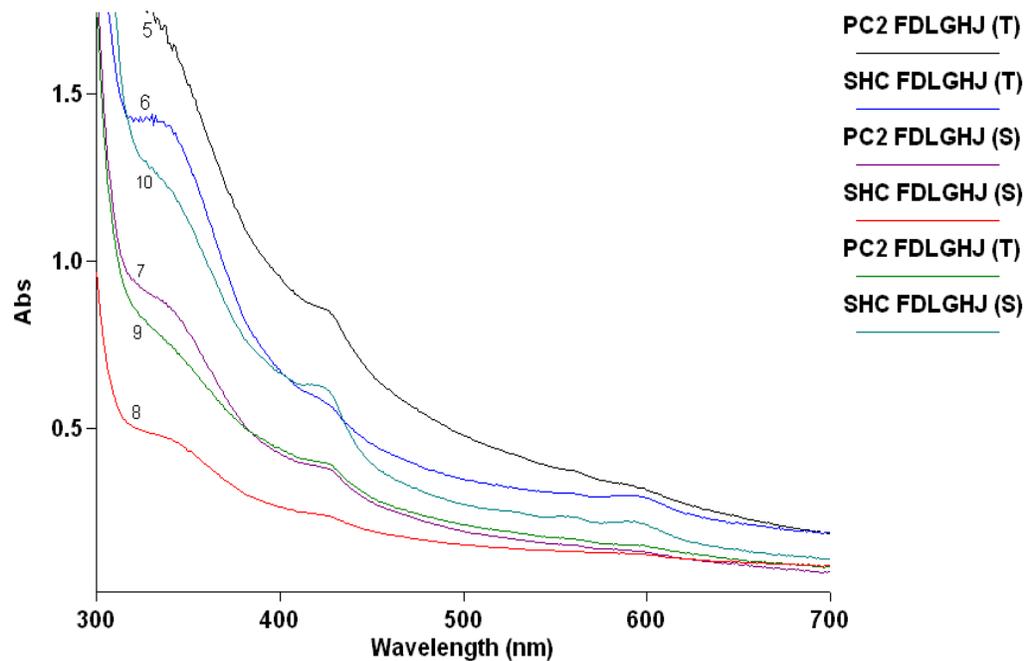
**Figure 5.11: UV-Vis spectra of NirF, D, L, G, H, J, incubated anaerobically with PC-2 and SHC with all cofactors.**

Overlaid spectra below show the UV-Vis scans of the resultant overnight incubations with PC-2 and SHC with isolated supernatants of overproduced NirF, D, L, G, H, J, proteins mixed together (T) and kept separately (S) prior to incubation. No difference is observed.



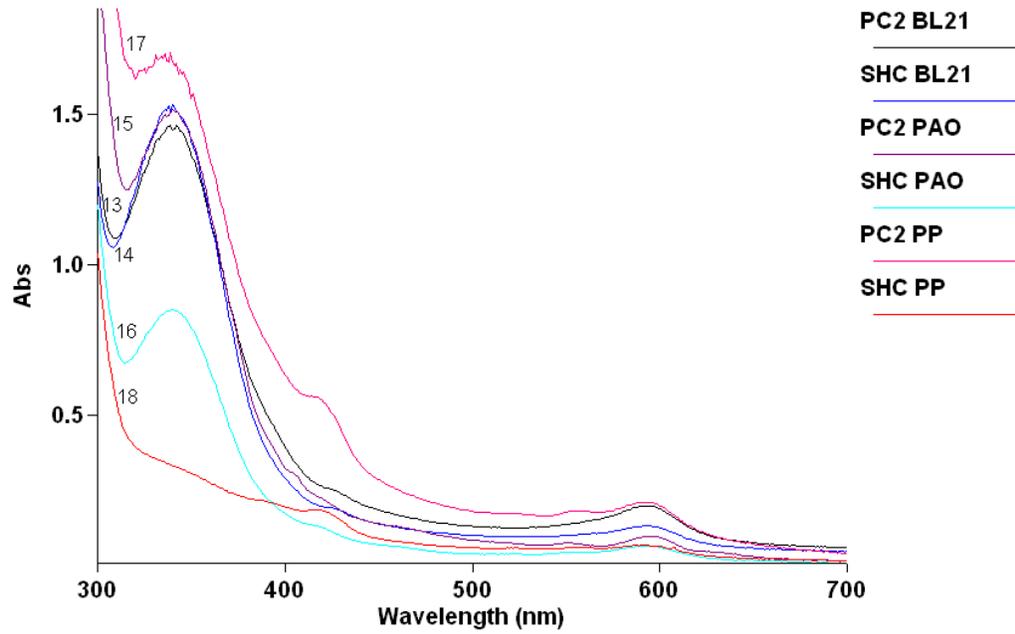
**Figure 5.12: UV-Vis spectra of NirF, D, L, G, H, J incubated anaerobically with PC-2 and SHC, cell extracts of *Ps. aeruginosa* and *Pa. pantrophus* and all cofactors.**

Overlaid UV-Vis scans of the resultant overnight incubations with PC-2 and SHC with isolated supernatants of overproduced NirF, D, L, G, H, J, proteins mixed together (T) and kept separately (S) prior to incubation, with *Ps. aeruginosa* and *Pa. pantrophus* cell extract and all cofactors. No difference is observed from the two methods.



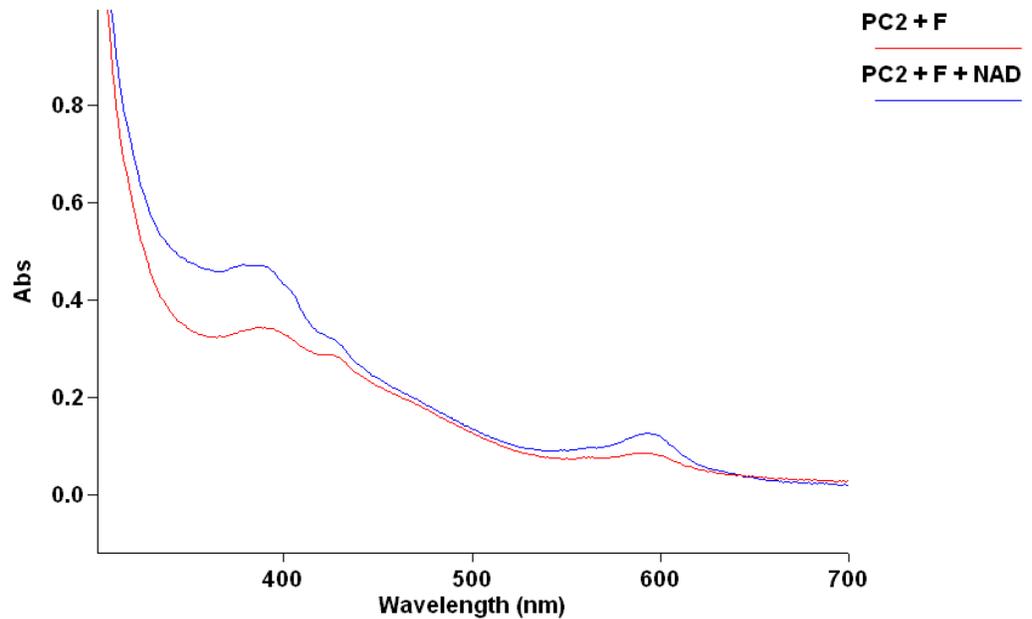
**Figure 5.13: UV-Vis spectra of cell extracts of *Ps. aeruginosa*, *E. coli* and *Pa. pantotrophus* incubated anaerobically with PC-2 and SHC and all cofactors (No Nir proteins).**

Overlaid UV-Vis scans of the resultant overnight incubations with PC-2 and SHC with all cofactors and isolated cell extract from *E. coli* and *Ps. aeruginosa* and *Pa. pantotrophus* as controls.

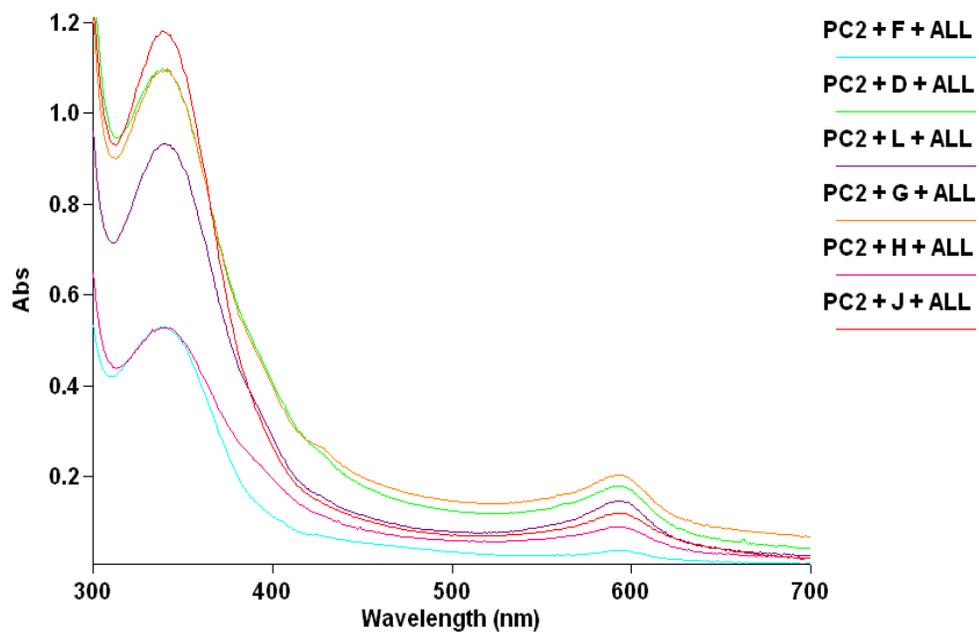


**Figure 5.14: UV-Vis spectra of anaerobic incubations with PC-2 using cell extracts of each overproduced Nir protein separately.**

**A** shows data from incubation of NirF with PC-2 with and without the addition of NAD<sup>+</sup>.

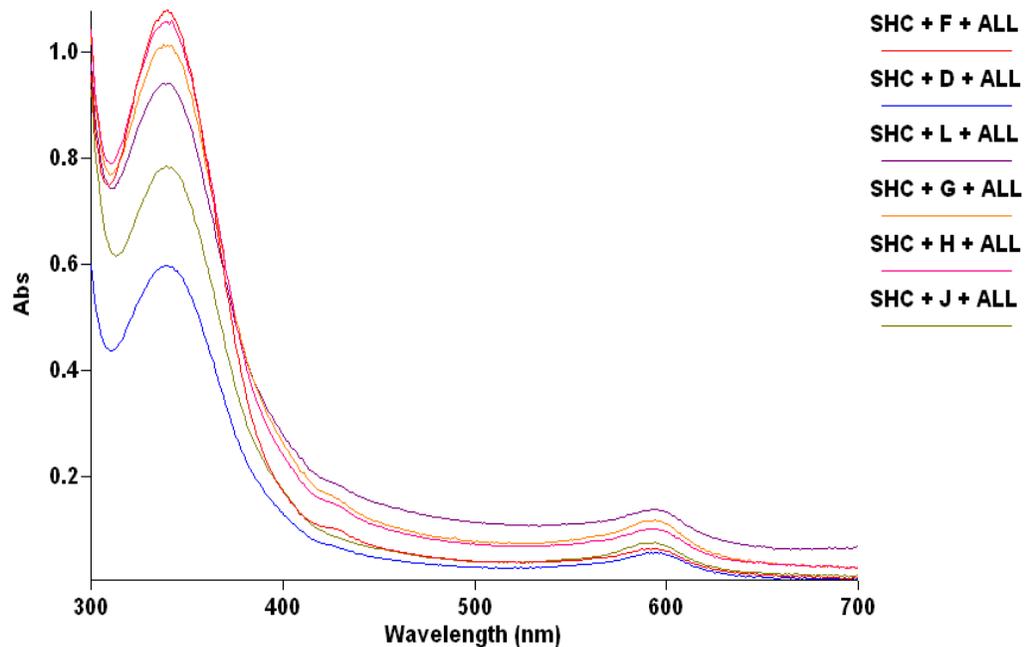


**B** shows the result of incubation with PC-2 of each Nir protein separately including all cofactors. No difference is observed between the proteins.



**Figure 5.15: UV-Vis spectra of anaerobic incubations with SHC using cell extracts of each overproduced Nir protein separately.**

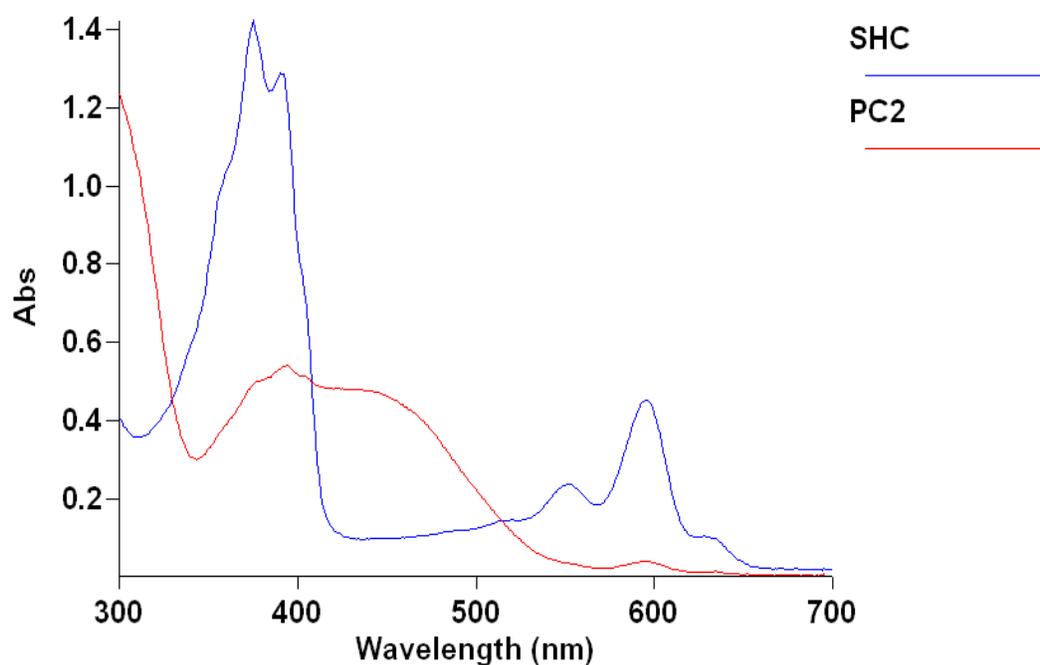
UV-Vis scans of the resultant overnight incubations with SHC of each Nir protein separately including all cofactors. No difference is observed between the proteins.



### 5.3.3. Results obtained from incubation of Nir proteins with PC-2 and SHC produced *in vitro*.

**Figure 5.16: UV-Vis spectra PC-2 and SHC produced from purified HemB, HemC, HemD, CobA ± SirC respectively ( $\pm\text{NAD}^+$ )**

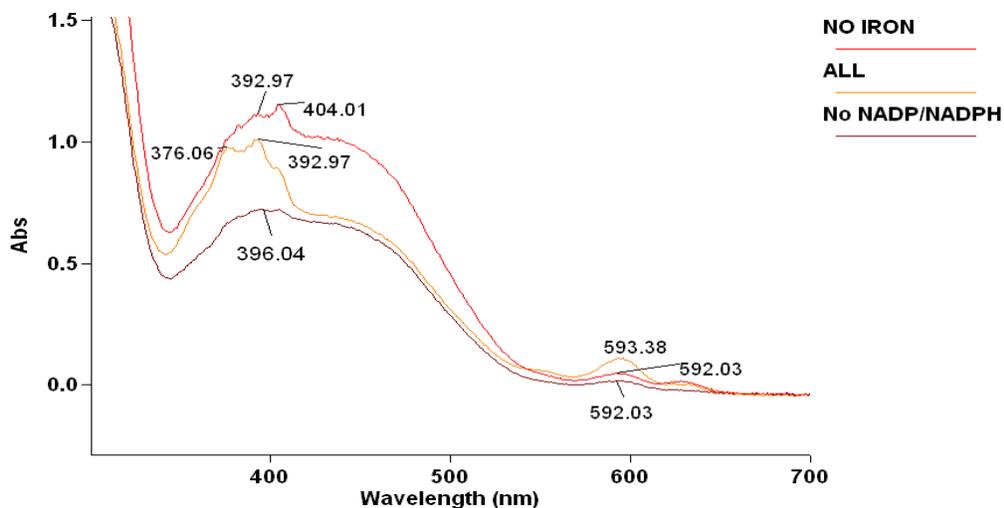
The resultant spectrum of PC-2 was diluted 1 in 10 and for SHC was diluted 1 in 20. Evidently this method of production allows for high levels of substrate and eliminates any background activity that may be seen in the spectra of PC-2 and SHC produced from soluble cell extracts.



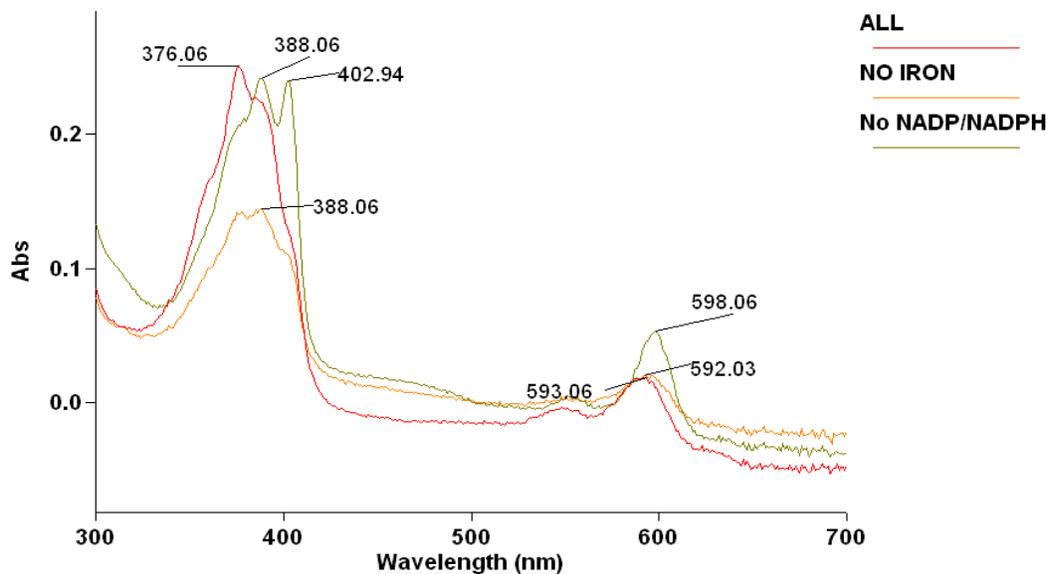
**Figure 5.17: UV-Vis spectra of compounds eluted at 1M NaCl from anaerobic incubation of PC-2 with NirFDLGHJ with varying cofactors (39-41).**

Fractions eluted at 1M, are shown in graphs **A** and **B**, for each incubation, one with all cofactors, the next with all cofactors except iron and one with all cofactors except NADP/NAPH

**A** shows the spectrum of fraction 1 eluted at 1M NaCl. This was yellow in colour and is PC-2, but in the process of oxidation to SHC.

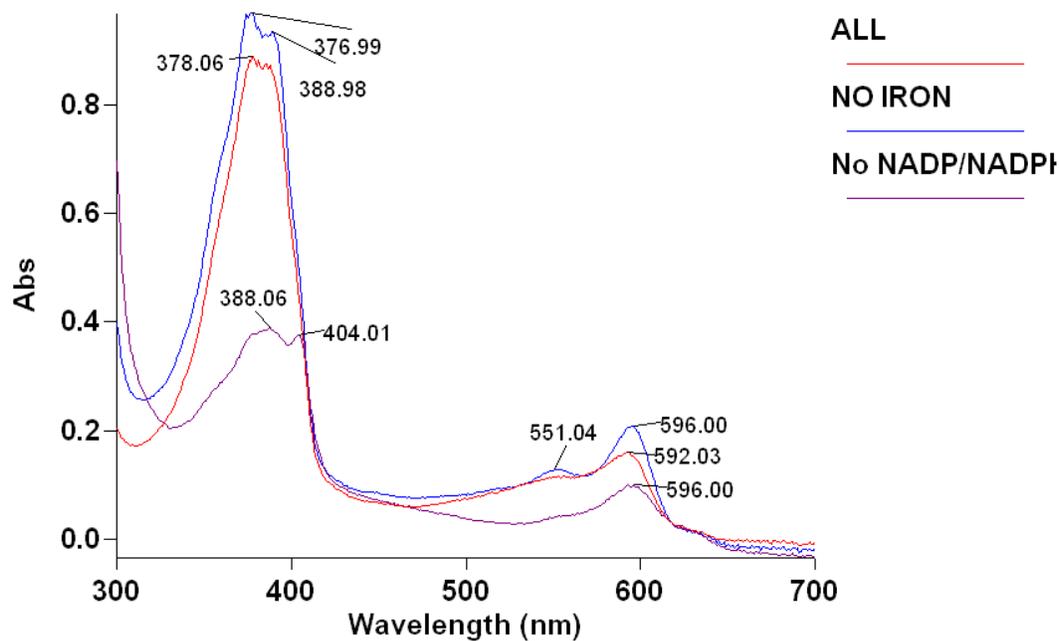


**B** is a pink fraction, eluted second, showing similarities to SHC, having absorbance maxima in the region of 376 nm and 590 nm. There was significantly less of this fraction than of that seen in graph A. Interestingly the absorbance maxima are in the region of 388 nm and 402 nm for the incubation without NADP/NADPH.



**Figure 5.18: UV-Vis spectra of compounds eluted at 1M NaCl from the anaerobic incubation of SHC with NirFDLGHJ and all cofactors (42-44).**

Fractions eluted at 1M are shown. Incubation of SHC and NirF, D, L, G, H, J with all cofactors, all cofactors except iron and with all cofactors except NADP/NAPH shows separation of a purple colour, with the characteristics of SHC.



#### 5.3.4. Discussion of data obtained from incubation of cell extracts containing overproduced Nir proteins with PC-2 and SHC produced *in situ* or *in vitro*.

Separation of the incubations as described in Table 5.2, showed no accumulation of haem  $d_1$  or any such new tetrapyrrole. Further analysis by the addition of *Ps. aeruginosa* and *Pa. pantrophus* cell extract to the incubation was performed to see if any factor present in these denitrifying organisms is required for the formation of haem  $d_1$  or its intermediates, since the proteins are being overproduced in *E. coli*. Other than the small differences expected between the organisms, shown in Figure 5.12, no difference was observed between the Nir proteins with addition of new cell extracts (5-12) from the controls (13-18: Figure 5.13). The analysis of the separation again revealed no accumulation of haem  $d_1$  or new tetrapyrroles suggesting the presence of *Ps. aeruginosa* and *Pa. pantrophus* cell extract is not a necessity. No significant difference is observed in the UV-Vis spectra when the Nir proteins are combined prior to sonication (T) or in the reaction mixture in the glove box (S) from the data obtained (1-4: Figure 5.11) again little difference is seen in comparison to the controls indicating this had no effect (13-18: Figure 5.13).

It was anticipated that the incubations performed would result in haem  $d_1$  or a new tetrapyrrole within the pathway from PC-2. This was not the case, suggesting that either the conditions are not exacting for haem  $d_1$  production or that a significant factor was missing. The added cofactors are based upon the requirements that may or may not occur in the formation of haem  $d_1$ . For example, if we have SHC as part of our pathway then previous work shows the dependence of dehydrogenase activity on  $\text{NAD}^+$ . As present it is unknown whether the pathway to haem  $d_1$  is via SHC but the regio-specificity of the methyl groups on the haem  $d_1$  macrocycle is that same as that of SHC. Thus, the activity of each Nir proteins towards PC-2 was assessed with the addition of  $\text{NAD}^+$  incubated overnight. The results shown in Figure 5.14A revealed no single Nir protein produced SHC from PC-2 in the presence of  $\text{NAD}^+$ , only NirF is shown as a representative spectra, as they showed no difference (19, 20); the incubation of each protein again with all the cofactors present with PC-2 similarly showed no SHC or any difference to each other as represented by Figure 5.14B (21-26).

Exactly the same reactions were performed this time with SHC instead of PC-2 only to show again that neither Nir protein was converting SHC into any newly derived tetrapyrrole (27-32: Figure 5.15). Even though two NirJ proteins from different organisms were present (*Ps. aeruginosa* and *Pa. pantrophus*) no change was observed. To rule out if buffer exchange of the cell extract was in some way having an effect the same incubations 1, 2, 5, 6, 9, 10 were repeated with out buffer exchange of the cell extracts (33-38, data not shown). Nothing was observed in the separation of the reaction mixtures again (data not shown) so this was not a factor.

It was noted that in combining the entire cell extracts a certain amount of background activity was evident in the spectra. To reduce this, and also to increase the amount of substrate available, PC-2 and SHC were synthesised from ALA by using purified proteins HemB, HemC, HemD, CobA ± SirC respectively. This method does not require filtration prior to use. The spectra obtained from PC-2 and SHC synthesised in this way is shown in Figure 5.16 and illustrates the high concentration of substrate available for use. Accordingly the incubations as detailed in Table 5.2 (39-44) were performed, with NirF, D, L, G, H, and J sonicated together.

The spectra obtained for the incubation of NirF, D, L, G, H, J, with a more concentrated amount of PC-2 and SHC substrate and thus less background activity showed no separation of haem  $d_1$ . Interestingly, separation of incubations with PC-2 isolated anaerobically (Figure 5.17B), show the presence of the same pink pigment observed in the expression of the operon or combination of *nir* genes in *E. coli* (Figure 5.6). This again showed some resemblance to SHC showing similar absorbance maxima in the region of 376 nm. Evidently there is a very small amount of this material due to the scale of the spectra. A lack of iron in the reaction mixture revealed less of this material, but the absence of NADP or NADPH was seen to shift the absorbance maxima to 388 nm and 402 nm. This may be suggestive of some dependency on iron for this material and exertion of some effect by the cofactors. The fluorescence of this pigment when illuminated under UV light was pink for this material and addition of strong acid turned it colourless whilst the addition of strong base turned it yellow. This pigment was however not observed in the expression of the operon in *Ps. aeruginosa* indicating that it may be confined to *E. coli* and thus not a true intermediate of the haem  $d_1$  pathway.

Incubations with SHC again (Figure 5.18), showed no isolation of any new tetrapyrrole or indeed haem  $d_1$ . Notably, the incubations with SHC and all the cofactors and without NADPH or NADP revealed the presence of a blue pigment exhibiting similarities to iron SHC (absorbance maxima at 412nm, data not shown), but since it was not observed in the incubation without iron, one may infer that one of the Nir proteins may be acting as a chelater of the exogenous iron to SHC and warrants further investigation. However, time was not a permitting factor and this investigation is confined to the future.

#### 5.4. Conclusions

This chapter explores the attempts to generate haem  $d_1$ . In doing so, it was hoped that this would also lead to the generation of intermediates in the haem  $d_1$  pathway. A number of methods were employed to synthesise these compounds both *in vivo* and *in vitro* under the assumption that a marked change in the spectral properties would be observed.

Primarily, bacteria were engineered to overexpress the haem  $d_1$  operon, *nirFDLGHJE* or *nirEFDLGHJ* for the generation of haem  $d_1$  *in vivo* in both *E. coli* and *Ps. aeruginosa*. Separation of the resultant cell extracts showed no accumulation of haem  $d_1$  in both organisms aerobically or anaerobically in *Ps. aeruginosa*. Further overexpression of *nirE*, *nirEF* *nirEFD*, *nirEFDL*, *nirEFDLG*, *nirEFDLGH* in *Ps. aeruginosa* did not exhibit or bind any substance at all when separated, and was observed to flow completely through the column. This was unexpected, as overexpression in *Ps. aeruginosa* was presumed to be more favourable for haem  $d_1$  or any intermediates. It is well documented that the expression of the *nir* genes is up regulated under anaerobic conditions and in particular with nitrate, at an optimal concentration of 25 mM (Carlson and Ingraham 1983; Filiatrault, Wagner et al. 2005). Such a result may be attributed to a higher requirement for substrates and a lower level of overproduction as observed in Chapter 4.

It was anticipated that since *in vivo* production of haem  $d_1$  was to no avail that the *in vitro* incubation of soluble cell extracts containing the overproduced Nir proteins with PC-2 and SHC would be more productive. However, incubations performed

with PC-2 and SHC produced *in situ*, with all the Nir proteins together and singularly, revealed no such presence of haem  $d_1$  or any new tetrapyrroles, respectively. It was notable that no Nir protein showed activity towards PC-2 in the presence of  $\text{NAD}^+$ , since the haem  $d_1$  macrocycle shares the same regio-specificity in its methyl groups as that of SHC. It is unknown if the pathway to haem  $d_1$  proceeds via SHC but the results obtained here indicate that it is unlikely as no Nir protein was able to convert PC-2 to SHC, and similarly no Nir protein showed activity towards SHC. This may be suggestive of a novel route for the synthesis of haem  $d_1$ , from PC-2 but not via SHC. Such a route has been observed for protohaem synthesis in *D. vulgaris*, whereby instead of decarboxylation of UIII, as with the well-characterised aerobic haem pathway, UIII is methylated to form PC-2. Synthesis then proceeds from PC-2 by decarboxylation and deacetylation to form coproporphyrinogen III, which is part of the aerobic pathway, that is then converted to protoporphyrinogen IX (Ishida, Yu et al. 1998). Therefore, it is likely that the biosynthesis of haem  $d_1$  occurs via PC-2, but thereafter a completely different pathway is employed.

Notably, overexpression of the multigene constructs in *E. coli* revealed the presence of a pink pigment with spectral similarities to SHC, with absorbance maxima in the region of 375 nm. This same pigment, though observed in small quantities, is observed when the Nir proteins are incubated with PC-2 at a high concentrations, but since it is not observed in *Ps. aeruginosa*, nor revealed when the proteins are singularly incubated with PC-2 or SHC, one may assume that this may be an artifactual result caused by non-exacting conditions in an organism that does not encode for these genes naturally. It may therefore be advantageous to check for activity towards PC-2 and SHC with the purified forms of the Nir proteins as this would eliminate any such pigments within the supernatant and allow for isolation of intermediates without interference from endogenous factors.

A blue pigment similar to that of iron SHC was observed when the Nir proteins were incubated with a high concentration of SHC and exogenous iron. This was not observed when iron was excluded from the reaction mixture and one may assume that the presence of one or more Nir proteins is able to chelate exogenous iron onto to SHC macrocycle. It would be highly desirable to check which of the Nir proteins or combinations thereof are able to chelate iron, or other metals, onto SHC, as this

would identify which enzyme is acting as a chelatase of the haem  $d_1$  pathway. Further investigations with metal free haem  $d_1$  would also be invaluable to check which Nir protein acts as a chelatase. Both experiments may allow differentiation of early or late chelation in the pathway.

Therefore, in summary, no conclusive evidence has been obtained that the *nir* operon, *nirFDLGHJE*, is able to synthesise haem  $d_1$  either *in vitro* or *in vivo*. Some interesting observations are evident, but these all require further detailed examination. No large significant changes were evident and those that were observed were on a small scale, these are less amenable and require further analysis by more improved and sensitive apparatus, such as HPLC combined with MS.

This chapter has provided some new and interesting data. It is the first concerted effort to engineer strains to produce haem  $d_1$  *in vivo* and further to synthesise intermediates *in vitro* by checking for activity towards well known tetrapyrrole precursors, PC-2 and SHC. It provides the framework for future studies on the synthesis of haem  $d_1$  and its intermediates and a valuable resource of multigene constructs to aid in its elucidation.

## CHAPTER 6

---

### **Conclusions**

## 6. Conclusions

The purpose of this PhD was to investigate the biosynthesis of a unique tetrapyrrole, haem  $d_1$ , whose pathway is not yet fully characterised. The elucidation of the pathway from UIII has posed a great challenge for over 20 years. This thesis has focused on many aspects of this pathway and explores many avenues to gain a deeper understanding of the biochemistry and molecular biology of such an important element of denitrification, and thus in nature.

The distinct lack of information regarding any prerequisite intermediates in this pathway is reflected in the similar lack of understanding as to how the aforementioned atypical substituents are introduced to the macrocycle. The introduction of oxygen on the macrocycle is of particular interest as this is achieved under anaerobic conditions. Again, the scarcity in the information is compounded with a lack of similarity of the proposed enzymes to those of known structure and function.

### **Confirmation that NirE is an active SUMT.**

In Chapter 3, it is shown for the first time that the previously proposed SUMT, NirE, is indeed able to catalyse the transformation of UIII to PC-2. Thus far, this had only been postulated on the basis of sequence similarity to the well-known SUMT enzymes, and had never been proven experimentally, until now. Mutational and complementation analysis (Palmedo, Seither et al. 1995; Kawasaki, Arai et al. 1997), has showed NirE is necessary for haem  $d_1$  production suggesting that flux through the haem  $d_1$  pathway was tightly regulated by NirE, as SUMTs are branch point enzymes displaying feedback inhibition to both substrate and product. NirE was purified to homogeneity and found to be a homodimer by gel filtration analysis. It also displays the typical characteristics attributed to SUMT enzymes, such as the formation of a red pigment, trimethylpyrrocorphin, which fluoresces when illuminated by UV light. NirE is unique though, in that it remains tightly bound to this porphyrinoid through purification, suggestive of a lack of inhibition of activity. This is confirmed by recent findings by Layer *et al*, 2009 (FEBS Journal, in press), whose work revealed NirE substrate inhibition occurs above [UIII] of  $17\mu\text{M}$ . This being much higher than CobA ( $2\mu\text{M}$ )(Blanche, Debussche et al. 1989), but is similar to that of CorA ( $20\mu\text{M}$ ) (Blanche, Robin et al. 1991). This is suggestive of the high requirement of NirE to generate sufficient levels of substrate for anaerobic

respiration, and thus no inhibition at low [UIII] facilitates greater flux through the pathway.

**Optimisation of overproduction and stabilisation revealed NirG as a tetramer and a monomer, and that NirD associated to NirG.**

Chapter 4 explores the challenges posed in the optimisation of overproduction of NirF, NirD, NirL, NirG, NirH, and NirJ. It is difficult to ascertain any information as to the reaction or reactions they may catalyse, as no similarity exists to proteins of known structure and function. A number of approaches were employed, whereby heterologous overproduction was successful, and a new system for homologous overexpression was formulated. However, homologous proteins were not overproduced to a sufficient level for further analysis and those overproduced heterologously were not stable. Introduction of all of these proteins into a stabilisation screen of various buffer conditions resulted in only NirG being suitably stabilised. Gel filtration analysis of NirG revealed for the first time, the presence of a tetramer as well as a monomer, revealing the preference of NirG to exist as part of a larger complex. It has long been proposed that NirD, NirL, NirG and NirH may exist as a complex. Therefore, it was predicted on the basis of the complementarity of NirD to NirG that co-overproduction of NirD, NirL, NirG and NirH with only NirD His-tagged would likely result in NirG being co-purified with NirD. Here we report for the first time that NirD does indeed bind NirG as they are co-purified together. In the same vein one would presume that NirL and NirH would be co-purified together also but requires further investigation.

No firm assignments can be made for the reactions these enzymes may catalyse. One can presume that NirDLGH performs the symmetrical changes to the macrocycle, such as decarboxylation at C12 and C18 and loss of propionic side chains at C3 and C8. The addition of oxygen to the macrocycle is likely to be performed by NirJ as similarity to HemN, an anaerobic oxidative decarboxylase, and radical SAM enzymes, suggest that NirJ is able to partake in the unique chemistry required to add oxygen on to the macrocycle. This leaves a putative assignment for NirF in ferrochelation. The elucidation of these possible functions is highly desirable as is the resolution of their tertiary structure. Structural information would be invaluable in determining function by comparing to proteins of similar topology. Furthermore, this would aid elucidation of the reactions and thus intermediates they may catalyse. Only NirG was suitably stable to be entered into

crystal trials, but analysis of crystals obtained showed them to be salt. Therefore further refinement is required to stabilise the Nir proteins to enter them into structural studies. It is noteworthy that changing the type of tag used to purify these proteins may indeed aid in stabilisation, but time did not permit this. It would be highly desirable to do this and thus obtain Nir proteins that are sufficiently stable to be entered in various structural and functional studies.

### **Attempts to generate haem $d_1$ and its intermediates *in vivo* and *in vitro* require further refinement**

The pathway from PC-2 to haem  $d_1$  has only one proposed intermediate, C800 (Youn, Liang et al. 2004), but it is unknown if it is indeed related in any way to the haem  $d_1$  pathway, either as a true intermediate or a derivative of an intermediate. Its ambiguity and unusual substituents, alongside the methods employed to connect this compound to the haem  $d_1$  pathway, leads to issue of the validity of such a material. Therefore, the isolation of a true intermediate of this pathway is highly sought-after. Chapter 5 explores the various approaches used to synthesise haem  $d_1$  and thus any intermediates *in vivo* and *in vitro*. *E. coli* and *Ps aeruginosa* were both engineered for the *in vivo* production of haem  $d_1$ . This was achieved by the Link and Lock method of cloning where the *nir* operon was successively cloned into a single plasmid, by utilising the ability of neoisoschizomers to generate the same cohesive ends which when ligated do not form a restrictable site. These sites were then used to clone each gene consecutively into the plasmid to form the operon. However, no production of haem  $d_1$  or any such new tetrapyrrole was observed in either organism by overexpression of the operon, or combinations of the *nir* genes. This is attributed to the low steady state concentrations of the enzymes, revealed in Chapter 4. Consequently, an *in vitro* approach was employed to optimise the level of the proposed Nir enzymes and thus, haem  $d_1$  or tetrapyrrole production.

In this approach, soluble cell extracts containing overproduced Nir proteins were checked for activity toward PC-2 and SHC, but again, no accumulation of haem  $d_1$  was seen and further no activity of any singularly incubated protein was observed towards PC-2 or SHC. Interestingly, a small amount of a pink pigment was observed in *E.coli* overexpressing *nirEFDLGHJ* and *nirFDLGHJE*. This pink pigment was also observed in the incubation of the cell extracts of all the Nir proteins towards PC-2. This pigment displays some similarity to SHC but since it

was not observed in *Ps. aeruginosa* or in response to incubation with SHC, one may presume that it is a result of overproduction of proteins not native to *E. coli*. It may therefore be advantageous to check for activity towards PC-2 and SHC with the purified forms of the Nir proteins as this would eliminate any such pigments within the supernatant and allow for isolation of intermediates without interference from endogenous factors.

The presence of a blue pigment was also seen in small quantity in the incubation of SHC with the Nir proteins and exogenous iron, but not in the absence of iron. This portrayed a spectrum similar to that of iron-SHC and may be indicative of iron chelation to SHC by one of the Nir proteins. It would be highly advantageous to investigate this further to check which of the Nir proteins or combinations thereof are able to chelate iron, or other metals, onto SHC, as this would identify which enzyme is acting as a chelatase of the haem  $d_1$  pathway. Investigations with metal free haem  $d_1$  would also be invaluable to check which Nir protein is able to chelate iron on to the haem  $d_1$  macrocycle. In this manner a candidate ferrochelatase may be identified.

It was surprising that neither method was able to form haem  $d_1$  or accumulate any intermediates. It was presumed that the overexpression of the operon or combinations of *nir* genes in *Ps. aeruginosa* would be particularly amenable, since they are being overexpressed in their native organism. Clearly, there is a factor missing or the conditions were not exacting since even haem  $d_1$  is not formed, as it was assumed that a large change would be evident in the UV-Vis spectral properties. Therefore, it would be beneficial to use methods more sensitive to low levels of tetrapyrroles such as HPLC coupled to MS to detect any smaller changes that may have not been previously evident. Such a method would allow small steady state levels of the tetrapyrrole intermediates to be detected.

It must also be considered that the *nir* operon is always preceded by *nirC* and is followed by *nirN*, both encoding for *c*-type cytochromes. It is presumed that NirC may be involved in maturation of the  $d_1$  haem and but studies by Hasegawa *et al*, 2001 found this was not the case showing NirC to be an electron donor to nitrite reductase, therefore this was not used in our investigations. Similarly, NirN was also omitted from our investigation but in retrospect it would be interesting to co-

express these with our *nir* operon to see if they exert any effect. Do they therefore play a role in haem  $d_1$  formation or stability or another effect?

The *Pa. pantotrophus nir* operon could be utilised in a similar fashion to the investigations performed in this thesis, as these may be more amenable to further investigation. Clearly, the most insoluble protein from *Ps aeruginosa* was NirJ. In stark contrast, NirJ from *Pa. pantotrophus* was largely soluble and stable. This may be a reflection of the ease of which the remaining Nir enzymes from *Pa. pantotrophus* may be used in the future and can be attributed to the higher level of haem  $d_1$  production in this organism. Therefore the level of the enzymes and consequently any intermediates would be higher in this organism and elucidation of the pathway made possible.

#### **Can we predict the route of haem $d_1$ synthesis on the basis of the well-characterised haem pathways?**

The synthesis of haem  $d_1$  necessitates methylation (C2, C7), decarboxylation (C12, C18), loss of propionates (C3, C8), oxidation (C3, C8), dehydrogenation for acrylate formation (C17), and ferrochelation. This thesis confirms the methylation of UIII by NirE to form PC-2. It is unknown if the pathway to haem  $d_1$  proceeds from PC-2 via SHC as the haem  $d_1$  macrocycle shares the same regio-specificity in its methyl groups as that of SHC (and PC-2). No Nir enzyme can be assigned this role, as there is no similarity to known PC-2 dehydrogenase enzymes. This thesis further illustrates that no Nir protein, or any combination of, showed activity towards PC-2 to generate SHC, in the presence of  $\text{NAD}^+$  and similarly no Nir protein showed activity towards SHC either. Therefore, on the basis of the data obtained within this study, one may presume that the synthesis of haem  $d_1$  proceeds via a completely separate and novel route, from PC-2 but not via SHC.

After consideration of the available information on the well characterised haem pathways and of the steps required to synthesise haem  $d_1$ , some parallel reactions are apparent and one can postulate a possible direction of synthesis for haem  $d_1$ . A similarity is seen to a novel route for protohaem biosynthesis observed in *D. vulgaris* (Ishida, Yu et al. 1998) and *M. barkeri* (Buchenau, Kahnt et al. 2006), where the genes for UIII synthesis are present but the late haem genes are absent, despite the presence of cytochromes. This novel pathway proceeds from UIII, where instead of decarboxylation of the acetate side chains of UIII, as with the well-

characterised aerobic haem pathway, UIII is methylated to form PC-2 (Ishida, Yu et al. 1998). This is parallel to the pathway for haem  $d_1$ , which also methylates UIII, forming PC-2 but thereafter it does not proceed via SHC, as with the alternative route for synthesising haem. The possible route for haem  $d_1$  biogenesis is discussed but one must be mindful of the scarcity of information available on the proposed gene products.

Synthesis of protohaem then proceeds from PC-2 by decarboxylation of the acetates at C12 and C18 to form 12,18 didecarboxyprecorrin-2. Since the synthesis of the haem  $d_1$  macrocycle also necessitates decarboxylation at C12 and C18 one may presume this would also be the route, shown in Figure 6.1. However, none of the proposed *nir* genes encodes for a decarboxylase enzyme. Although there exists some similarity of NirF to CobT, a decarboxylase of the cobalamin pathway, the assignment of a defined function to the Nir proteins remains vague.

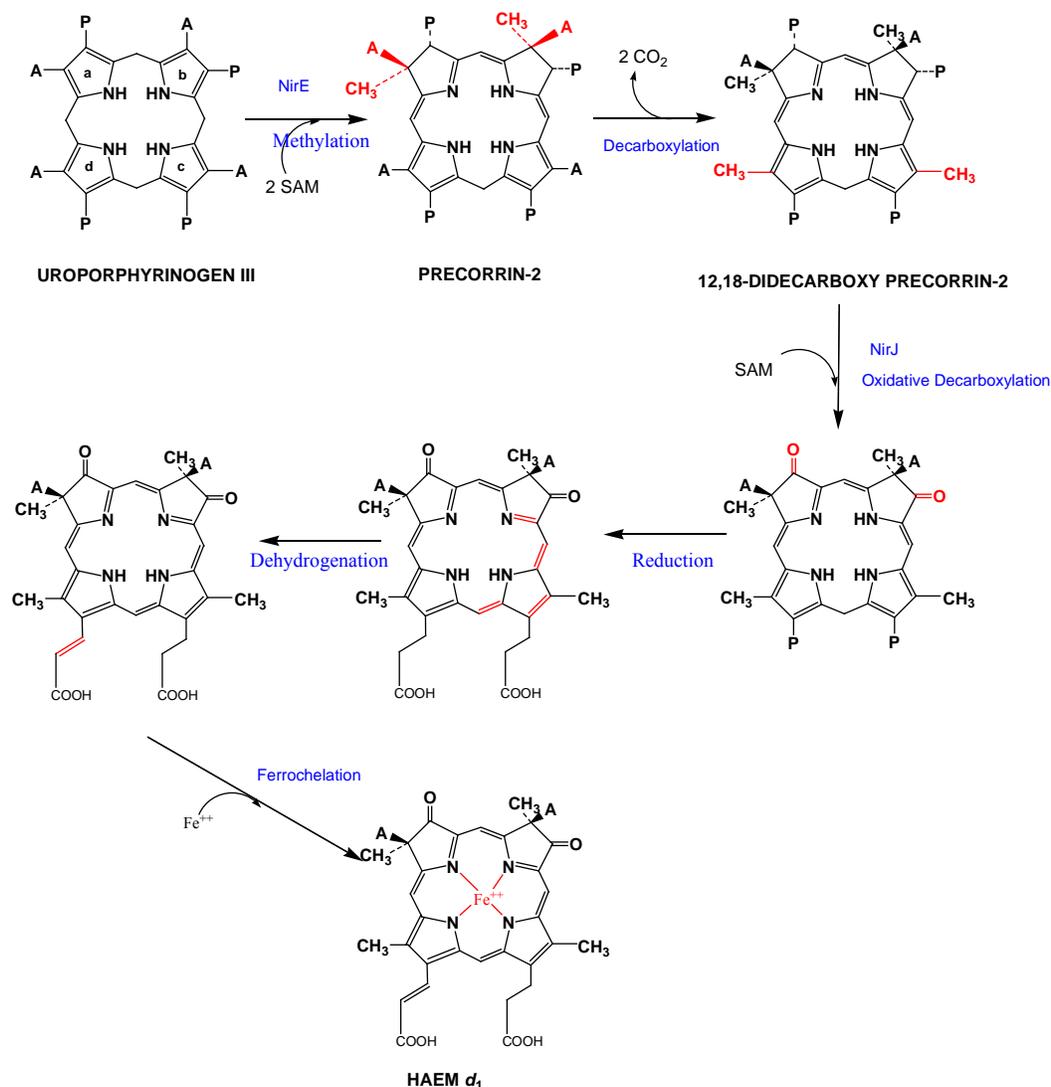
After decarboxylation at C12 and C18, the elimination of the acetate groups at C2 and C7 forms coproporphyrinogen III (CPIII), returning to the more conventional route for haem synthesis. CPIII is then oxidatively decarboxylated to protoporphyrinogen IX (PP'gen IX), by HemN, which converts the propionate substituents at C3 and C8 to vinyl groups. This is achieved by the unique radical SAM chemistry that the Fe-S moiety bestows upon the enzyme. In Chapter 4, the evident similarity of HemN to NirJ can allow the inference that NirJ incorporates oxygen on to the haem  $d_1$  macrocycle in a similar fashion to HemN, and it is likely to be performed after decarboxylation at C12 and C18, and removal of the propionates at C3 and C8, shown in Figure 6.1. NirJ may of course be able to mediate the transformation of these propionates to oxygen on the macrocycle at C3 and C8, similar to HemN, but the chemistry is different in that HemN forms vinyl groups not the keto groups of haem  $d_1$ . It must also be stated that the C3 and C8 keto groups are derived from the carboxylic acid group of the propionic acids side chains of the macrocycle (Matthews 1993) and not from another source, therefore it is likely that NirJ performs this transformation of the macrocycle.

The pathway to protohaem is completed by the oxidation of PP'gen IX to protoporphyrin IX and chelation of iron. In the case of haem  $d_1$ , the dehydrogenation of the C17 propionate to form the acrylate substituent and ferrochelation are required. Again, no Nir protein shows any similarity to proteins

that may undertake these reactions. One may return to the initial inferences, that NirDLGH may perform decarboxylation at C12 and C18 and remove the propionates at C3 and C8, thereby facilitating NirJ to oxidise the macrocycle at C3 and C8, and NirF chelates iron. However, again, this theory also does not explain the formation an acrylate. Therefore, it can be presumed that one or more of the Nir proteins is multifunctional and is able to afford more than one of these transformations to the haem  $d_1$  macrocycle, or in fact are part of a larger multimeric complex able to partake in a number of reactions together with the substrate still bound.

### Figure 6.1: Schematic representation of the putative pathway to haem $d_1$

The possible pathway to haem  $d_1$  is shown here and discussed in detail in the text. It does not proceed via SHC, but from PC-2 being similar to a novel haem pathway.



In summary, the pathway for haem  $d_1$  proposed in this thesis and presented in Figure 6.1, is that of methylation of UIII at C2 and C7 to yield PC-2 by NirE. PC-2 may then be decarboxylated at C12 and C18 to form 12,18 didecarboxyprecorrin-2, as found in the alternative haem pathway. It is unknown if decarboxylation is performed by NirF or a NirDLGH complex, but similarity of NirF to CobT lends to the possibility of NirF performing this reaction. The removal of the propionates at C3 and C8, with subsequent anaerobic oxidation of the macrocycle is, expected to be facilitated by NirJ. It is unknown how further saturation of the macrocycle and ferrochelation is afforded as no single Nir enzyme has been identified to partake in these transformations. It may indeed be performed by the multimeric NirDLGH, bestowing a number of roles partaking in the transformation of the preceding intermediates to haem  $d_1$ . It is logical to presume that ferrochelation is the last step of the pathway, as found in the haem pathways characterised thus far, but again no assignment can be made as no similarity exists to known chelatasases. This has been summarised in Table 6.1.

**Table 6.1: Summary of possible roles of the Nir Proteins**

A table summarising the possible reactions each of the Nir proteins are predicted to catalyse on the basis of the inferences made in this thesis

<b>Protein</b>	<b>Reactions catalysed /Possible role</b>
NirF	Decarboxylation of PC-2 to 12,18 -didecarboxy precorrin-2
NirD NirL NirG NirH	Multimeric complex performing saturation of macrocycle and ferrochelation
NirJ	Anearobic decarboxylation of 12,18 -didecarboxy precorrin-2
NirE	SUMT- methylation of UIII to PC-2

On the basis of the inferences made herein, it is apparent that all NirF, NirD, NirL, NirG, NirH, NirJ, NirE are required in the biogenesis of haem  $d_1$ , each presumably imparting a specific role and thus transformation afforded to the macrocycle. This may be as a single enzyme or as part of a larger multimeric complex and is in

conjunction with previous genetic studies demonstrating their necessity. In contrast, the proposed haem  $d_1$  pathway in *Heliobacter* and the proposed intermediate, C800, both do not withstand their theories. In the former a wholly dissimilar pathway for haem  $d_1$  is proposed based entirely on bioinformatic analysis, in an organism that is not documented to require or form haem  $d_1$ . Again, not all of the transformations are accounted for and further, the presence of some Nir proteins is excluded from the pathway and replaced by CysG. This would explain inclusion of SHC in their pathway, but the data represented in Chapter 4, clearly shows no Nir enzyme has activity towards PC-2 to form SHC. The proposed pathway appears to be a confusion of the alternative haem and sirohaem pathways, therefore it holds no bearing on haem  $d_1$  biosynthesis. In the latter, the formation of such an unusual intermediate is unlikely since the energy required to make haem  $d_1$  in this manner would not likely be favourable for anaerobic respiration, where a high demand for haem  $d_1$  occurs.

### **Epilogue**

Despite the many different attempts made in this thesis to gain a deeper understanding of the mechanism of such a unique pathway, the biosynthesis of haem  $d_1$  still remains uncharacterised after 20 years of research. This project has focused on many aspects of this pathway, from the investigation of the proteins encoded by operon, *nirF*, *nirD*, *nirL*, *nirG*, *nirH*, *nirJ* and *nirE* in *Ps. aeruginosa*, to the possibility of generating haem  $d_1$  or any intermediates *in vivo* and *in vitro*. Whilst no firm assignments can be made for each protein some interesting results were obtained and similarly, the lack of activity towards PC-2 or SHC further lend to the hypothesis that haem  $d_1$  follows a completely separate and unique pathway. A possible route for the biogenesis of haem  $d_1$  has been proposed on the basis of possible common mechanisms for the synthesis of protohaem. Evidently, further investigation is required to elucidate the pathway for the synthesis of haem  $d_1$ .

Therefore, this thesis documents the first concerted effort to study the haem  $d_1$  pathway in this manner and provides the framework for future studies on this molecule and its proposed enzymes. Consequently, it offers a valuable resource of single or multigene plasmid constructs and therefore potentially stable proteins and envisages a positive future for elucidation of the pathway for the future of haem  $d_1$  research, which is clearly a highly desirable, but challenging prospect.

## REFERENCES

---

- Arai, H., Y. Igarashi, et al. (1994). "Structure and ANR-dependent transcription of the nir genes for denitrification from *Pseudomonas aeruginosa*." Biosci Biotechnol Biochem **58**(7): 1286-91.
- Arai, H., Y. Igarashi, et al. (1995). "Expression of the nir and nor genes for denitrification of *Pseudomonas aeruginosa* requires a novel CRP/FNR-related transcriptional regulator, DNR, in addition to ANR." FEBS Lett **371**(1): 73-6.
- Arai, H., T. Kodama, et al. (1999). "Effect of nitrogen oxides on expression of the nir and nor genes for denitrification in *Pseudomonas aeruginosa*." FEMS Microbiol Lett **170**(1): 19-24.
- Averill, B. A. (1996). "Dissimilatory Nitrite and Nitric Oxide Reductases." Chem Rev **96**(7): 2951-2964.
- Baker, S. C., S. J. Ferguson, et al. (1998). "Molecular genetics of the genus *Paracoccus*: metabolically versatile bacteria with bioenergetic flexibility." Microbiol Mol Biol Rev **62**(4): 1046-78.
- Baker, S. C., N. F. Saunders, et al. (1997). "Cytochrome cd1 structure: unusual haem environments in a nitrite reductase and analysis of factors contributing to beta-propeller folds." J Mol Biol **269**(3): 440-55.
- Beale, S. I. W., J.D. (1990). Tetrapyrrole Metabolism in photosynthetic organisms: Biosynthesis of hemes and chlorophylls, McGraw-Hill, New York.
- Bell, L. C., D. J. Richardson, et al. (1990). "Periplasmic and membrane-bound respiratory nitrate reductases in *Thiosphaera pantotropha*. The periplasmic enzyme catalyzes the first step in aerobic denitrification." FEBS Lett **265**(1-2): 85-7.
- Berks, B. C., S. J. Ferguson, et al. (1995b). "Enzymes and associated electron transport systems that catalyse the respiratory reduction of nitrogen oxides and oxyanions." Biochim Biophys Acta **1232**(3): 97-173.
- Berks, B. C., M. D. Page, et al. (1995). "Sequence analysis of subunits of the membrane-bound nitrate reductase from a denitrifying bacterium: the integral membrane subunit provides a prototype for the dihaem electron-carrying arm of a redox loop." Mol Microbiol **15**(2): 319-31.
- Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Thibaut, D., Vuilhorgne, M., Leepe, F.J., Battersby, A.R. (1995). "Vitamin B12 - how the problem of its biosynthesis was solved." Angewandte Chemie Int Ed **34**: 383-411.
- Blanche, F., L. Debussche, et al. (1989). "Purification and characterization of S-adenosyl-L-methionine: uroporphyrinogen III methyltransferase from *Pseudomonas denitrificans*." J Bacteriol **171**(8): 4222-31.
- Blanche, F., C. Robin, et al. (1991). "Purification, characterization, and molecular cloning of S-adenosyl-L-methionine: uroporphyrinogen III methyltransferase from *Methanobacterium ivanovii*." J Bacteriol **173**(15): 4637-45.
- Braun, C. and W. G. Zumft (1992). "The structural genes of the nitric oxide reductase complex from *Pseudomonas stutzeri* are part of a 30-kilobase gene cluster for denitrification." J Bacteriol **174**(7): 2394-7.
- Buchenau, B., J. Kahnt, et al. (2006). "Heme biosynthesis in *Methanosarcina barkeri* via a pathway involving two methylation reactions." J Bacteriol **188**(24): 8666-8.
- Carlson, C. A., L. P. Ferguson, et al. (1982). "Properties of dissimilatory nitrate reductase purified from the denitrifier *Pseudomonas aeruginosa*." J Bacteriol **151**(1): 162-71.

- Carlson, C. A. and J. L. Ingraham (1983). "Comparison of denitrification by *Pseudomonas stutzeri*, *Pseudomonas aeruginosa*, and *Paracoccus denitrificans*." *Appl Environ Microbiol* **45**(4): 1247-53.
- Chang, C. K. (1985). "On the structure of heme d1. An isobacteriochlorin derivative as the prosthetic group of dissimilatory nitrite reductase?" *J Biol Chem* **260**(17): 9520-2.
- Chang, C. K. (1994). "Haem d1 and other haem cofactors from bacteria." *Ciba Found Symp* **180**: 228-38; discussion 238-46.
- Chang, C. K., L. K. Hanson, et al. (1981). "pi cation radicals of ferrous and free base isobacteriochlorins: Models for siroheme and sirohydrochlorin." *Proc Natl Acad Sci U S A* **78**(5): 2652-2656.
- Chang, C. K., R. Timkovich, et al. (1986b). "Evidence that heme d1 is a 1,3-porphyrindione." *Biochemistry* **25**(26): 8447-53.
- Chang, C. K. and W. Wu (1986). "The porphinedione structure of heme d1. Synthesis and spectral properties of model compounds of the prosthetic group of dissimilatory nitrite reductase." *J Biol Chem* **261**(19): 8593-6.
- Cheesman, M. R., S. J. Ferguson, et al. (1997). "Two enzymes with a common function but different heme ligands in the forms as isolated. Optical and magnetic properties of the heme groups in the oxidized forms of nitrite reductase, cytochrome cd1, from *Pseudomonas stutzeri* and *Thiosphaera pantotropha*." *Biochemistry* **36**(51): 16267-76.
- Chuanchuen, R., C. T. Narasaki, et al. (2002). "Benchtop and microcentrifuge preparation of *Pseudomonas aeruginosa* competent cells." *Biotechniques* **33**(4): 760, 762-3.
- Coyne, M. S., A. Arunakumari, et al. (1990). "Localization of the cytochrome cd1 and copper nitrite reductases in denitrifying bacteria." *J Bacteriol* **172**(5): 2558-62.
- Crane, B. R., L. M. Siegel, et al. (1995). "Sulfite reductase structure at 1.6 Å: evolution and catalysis for reduction of inorganic anions." *Science* **270**(5233): 59-67.
- Cronin, C. N. and W. S. McIntire (1999). "pUCP-Nco and pUCP-Nde: *Escherichia-Pseudomonas* shuttle vectors for recombinant protein expression in *Pseudomonas*." *Anal Biochem* **272**(1): 112-5.
- Crouzet, J., L. Cauchois, et al. (1990). "Nucleotide sequence of a *Pseudomonas denitrificans* 5.4-kilobase DNA fragment containing five cob genes and identification of structural genes encoding S-adenosyl-L-methionine: uroporphyrinogen III methyltransferase and cobyrinic acid a,c-diamide synthase." *J Bacteriol* **172**(10): 5968-79.
- Cutruzzola, F., K. Brown, et al. (2001). "The nitrite reductase from *Pseudomonas aeruginosa*: essential role of two active-site histidines in the catalytic and structural properties." *Proc Natl Acad Sci U S A* **98**(5): 2232-7.
- de Boer, A. P., W. N. Reijnders, et al. (1994). "Isolation, sequencing and mutational analysis of a gene cluster involved in nitrite reduction in *Paracoccus denitrificans*." *Antonie Van Leeuwenhoek* **66**(1-3): 111-27.
- de Boer, A. P., J. van der Oost, et al. (1996). "Mutational analysis of the nor gene cluster which encodes nitric-oxide reductase from *Paracoccus denitrificans*." *Eur J Biochem* **242**(3): 592-600.
- Debussche, L., D. Thibaut, et al. (1993). "Biosynthesis of the corrin macrocycle of coenzyme B12 in *Pseudomonas denitrificans*." *J Bacteriol* **175**(22): 7430-40.
- Dunn, N. W. and B. W. Holloway (1971). "Pleiotrophy of p-fluorophenylalanine-resistant and antibiotic hypersensitive mutants of *Pseudomonas aeruginosa*." *Genet Res* **18**(2): 185-97.
- Ermler, U., W. Grabarse, et al. (1997). "Crystal structure of methyl-coenzyme M reductase: the key enzyme of biological methane formation." *Science* **278**(5342): 1457-62.
- Fan, J., D. Wang, et al. (2006). "Maize uroporphyrinogen III methyltransferase: overexpression of the functional gene fragments in *Escherichia coli* and one-step purification." *Protein Expr Purif* **46**(1): 40-6.

- Ferguson, S. J. (2009). The role of Heme d1 in Denitrification. Tetrapyrroles: Birth, Life and Death., Landes Bioscience.
- Filiatrault, M. J., V. E. Wagner, et al. (2005). "Effect of anaerobiosis and nitrate on gene expression in *Pseudomonas aeruginosa*." Infect Immun **73**(6): 3764-72.
- Frank, S., A. A. Brindley, et al. (2005). "Anaerobic synthesis of vitamin B12: characterization of the early steps in the pathway." Biochem Soc Trans **33**(Pt 4): 811-4.
- Frank, S., E. Deery, et al. (2007). "Elucidation of substrate specificity in the cobalamin (vitamin B12) biosynthetic methyltransferases. Structure and function of the C20 methyltransferase (CbiL) from *Methanothermobacter thermoautotrophicus*." J Biol Chem **282**(33): 23957-69.
- Fulop, V., J. W. Moir, et al. (1995). "The anatomy of a bifunctional enzyme: structural basis for reduction of oxygen to water and synthesis of nitric oxide by cytochrome cd1." Cell **81**(3): 369-77.
- George, S. J., J. W. Allen, et al. (2000). "Time-resolved infrared spectroscopy reveals a stable ferric heme-NO intermediate in the reaction of *Paracoccus pantotrophus* cytochrome cd1 nitrite reductase with nitrite." J Biol Chem **275**(43): 33231-7.
- Gudat, J. C., J. Singh, et al. (1973). "Cytochrome oxidase from *Pseudomonas aeruginosa*. I. Purification and some properties." Biochim Biophys Acta **292**(2): 376-90.
- Hanahan, D. (1983). "Studies on transformation of *Escherichia coli* with plasmids." J Mol Biol **166**(4): 557-80.
- Hanahan, D. (1985). DNA Cloning, IRL Press Ltd. 109 - 135.
- Hansen, J., M. Muldbjerg, et al. (1997). "Siroheme biosynthesis in *Saccharomyces cerevisiae* requires the products of both the MET1 and MET8 genes." FEBS Lett **401**(1): 20-4.
- Harris III, W. f., Burkhalter, W.L., Timkovich, R (1993). "Enhancement of bacterial porphyrin biosynthesis by exogenous aminolevulinic acid and isomer specificity." Bioorg Chem(21): 209-220.
- Hartig, E. and W. G. Zumft (1999). "Kinetics of nirS expression (cytochrome cd1 nitrite reductase) in *Pseudomonas stutzeri* during the transition from aerobic respiration to denitrification: evidence for a denitrification-specific nitrate- and nitrite-responsive regulatory system." J Bacteriol **181**(1): 161-6.
- Hennig, M., B. Grimm, et al. (1997). "Crystal structure of glutamate-1-semialdehyde aminomutase: an alpha2-dimeric vitamin B6-dependent enzyme with asymmetry in structure and active site reactivity." Proc Natl Acad Sci U S A **94**(10): 4866-71.
- Hernandez-Ramirez, G., S. M. Brouder, et al. (2009). "Greenhouse gas fluxes in an eastern Corn Belt soil: weather, nitrogen source, and rotation." J Environ Qual **38**(3): 841-54.
- Hill, K. E. and D. C. Wharton (1978). "Reconstitution of the apoenzyme of cytochrome oxidase from *Pseudomonas aeruginosa* with heme d1 and other heme groups." J Biol Chem **253**(2): 489-95.
- Hille, R. (1996). "The Mononuclear Molybdenum Enzymes." Chem Rev **96**(7): 2757-2816.
- Hodgkin, D. G., J. Pickworth, et al. (1955). "The crystal structure of the hexacarboxylic acid derived from B12 and the molecular structure of the vitamin." Nature **176**(4477): 325-8.
- Holloway, B. W., U. Romling, et al. (1994). "Genomic mapping of *Pseudomonas aeruginosa* PAO." Microbiology **140** (Pt 11): 2907-29.
- Horio, T., T. Higashi, et al. (1961). "Purification and properties of cytochrome oxidase from *Pseudomonas aeruginosa*." J Biol Chem **236**: 944-51.
- Huang, D. D. and W. Y. Wang (1986). "Chlorophyll biosynthesis in *Chlamydomonas* starts with the formation of glutamyl-tRNA." J Biol Chem **261**(29): 13451-5.
- Hungerer, C., B. Troup, et al. (1995a). "Regulation of the hemA gene during 5-aminolevulinic acid formation in *Pseudomonas aeruginosa*." J Bacteriol **177**(6): 1435-43.

- Hungerer, C., B. Troup, et al. (1995b). "Cloning, mapping and characterization of the *Pseudomonas aeruginosa* hemL gene." Mol Gen Genet **248**(3): 375-80.
- Ishida, T., L. Yu, et al. (1998). "A primitive pathway of porphyrin biosynthesis and enzymology in *Desulfovibrio vulgaris*." Proc Natl Acad Sci U S A **95**(9): 4853-8.
- Johansson, P. and L. Hederstedt (1999). "Organization of genes for tetrapyrrole biosynthesis in gram--positive bacteria." Microbiology **145 (Pt 3)**: 529-38.
- Jones, M. C., J. M. Jenkins, et al. (1994). "Cloning and characterisation of genes for tetrapyrrole biosynthesis from the cyanobacterium *Anacystis nidulans* R2." Plant Mol Biol **24**(3): 435-48.
- Jungst, A., C. Braun, et al. (1991). "Close linkage in *Pseudomonas stutzeri* of the structural genes for respiratory nitrite reductase and nitrous oxide reductase, and other essential genes for denitrification." Mol Gen Genet **225**(2): 241-8.
- Kannangara, C. G., S. P. Gough, et al. (1988). "tRNA(Glu) as a cofactor in delta-aminolevulinic acid biosynthesis: steps that regulate chlorophyll synthesis." Trends Biochem Sci **13**(4): 139-43.
- Kawasaki, S., H. Arai, et al. (1995). "Sequencing and characterization of the downstream region of the genes encoding nitrite reductase and cytochrome c-551 (*nirSM*) from *Pseudomonas aeruginosa*: identification of the gene necessary for biosynthesis of heme d1." Gene **167**(1-2): 87-91.
- Kawasaki, S., H. Arai, et al. (1997). "Gene cluster for dissimilatory nitrite reductase (*nir*) from *Pseudomonas aeruginosa*: sequencing and identification of a locus for heme d1 biosynthesis." J Bacteriol **179**(1): 235-42.
- Korner, H. and F. Mayer (1992). "Periplasmic location of nitrous oxide reductase and its apofrom in denitrifying *Pseudomonas stutzeri*." Arch Microbiol **157**(3): 218-22.
- Kredich, N. M. (1987). "In *E. coli* and *S. typhimurium* cellular and molecular biology." American society of Microbiology: 419-428.
- Kuronen, T. and N. Ellfolk (1972). "A new purification procedure and molecular properties of *Pseudomonas* cytochrome oxidase." Biochim Biophys Acta **275**(3): 308-18.
- Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-5.
- Lashof, D. A., Ahuja, D.R. (1990). "Relative contributions of greenhouse gas emissions to global warming." Nature (3): 529-531.
- Layer, G., E. Kervio, et al. (2005). "Structural and functional comparison of HemN to other radical SAM enzymes." Biol Chem **386**(10): 971-80.
- Layne (1957). "Spectrophotometric and Turbidimetric Methods for Measuring Proteins." Methods in Enzymology **3**: 447-455.
- Leech, H. K., E. Raux-Deery, et al. (2002). "Production of cobalamin and sirohaem in *Bacillus megaterium*: an investigation into the role of the branchpoint chelatases sirohydrochlorin ferrochelatase (*SirB*) and sirohydrochlorin cobalt chelatase (*CbiX*)." Biochem Soc Trans **30**(4): 610-3.
- Leustek, T., M. Smith, et al. (1997). "Siroheme biosynthesis in higher plants. Analysis of an S-adenosyl-L-methionine-dependent uroporphyrinogen III methyltransferase from *Arabidopsis thaliana*." J Biol Chem **272**(5): 2744-52.
- Matthews, J. C. T., R (1993). "Biosynthetic origins of the carbon skeleton of heme d1." bioorganic chemistry **21**: 71-82.
- McGoldrick, H. M., C. A. Roessner, et al. (2005). "Identification and characterization of a novel vitamin B12 (cobalamin) biosynthetic enzyme (*CobZ*) from *Rhodobacter capsulatus*, containing flavin, heme, and Fe-S cofactors." J Biol Chem **280**(2): 1086-94.
- McPherson, A., Jr. (1976). "The growth and preliminary investigation of protein and nucleic acid crystals for X-ray diffraction analysis." Methods Biochem Anal **23**(0): 249-345.

- Micklefield, J., Beckmann, M, Mackman, R.L, Block M.H, Leeper, F.J, Battersby, A.R (1993). "A novel stereoselective synthesis of the macrocycle of haem d1 establishes its absolute configuration as 2*R*, 7*R*." Chem Soc, Chem Comm: 275-577.
- Micklefield, J., Beckmann, M, Mackman, R.L, Block M.H, Leeper, F.J, Battersby, A.R (1997). "Haem d1: stereoselective synthesis of the macrocycle to establish its absolute configuration as 2*R*, 7*R*." J. Chem. Soc., Perkin Trans (1).
- Murphy, M. J., L. M. Siegel, et al. (1974). "Siroheme: a new prosthetic group participating in six-electron reduction reactions catalyzed by both sulfite and nitrite reductases." Proc Natl Acad Sci U S A **71**(3): 612-6.
- Nurizzo, D., F. Cutruzzola, et al. (1998). "Conformational changes occurring upon reduction and NO binding in nitrite reductase from *Pseudomonas aeruginosa*." Biochemistry **37**(40): 13987-96.
- Nurizzo, D., F. Cutruzzola, et al. (1999). "Does the reduction of c heme trigger the conformational change of crystalline nitrite reductase?" J Biol Chem **274**(21): 14997-5004.
- Nurizzo, D., M. C. Silvestrini, et al. (1997). "N-terminal arm exchange is observed in the 2.15 Å crystal structure of oxidized nitrite reductase from *Pseudomonas aeruginosa*." Structure **5**(9): 1157-71.
- Page, M. D. and S. J. Ferguson (1994). "Differential reduction in soluble and membrane-bound c-type cytochrome contents in a *Paracoccus denitrificans* mutant partially deficient in 5-aminolevulinic synthase activity." J Bacteriol **176**(19): 5919-28.
- Palmedo, G., P. Seither, et al. (1995). "Resolution of the *nirD* locus for heme d1 synthesis of cytochrome cd1 (respiratory nitrite reductase) from *Pseudomonas stutzeri*." Eur J Biochem **232**(3): 737-46.
- Peakman, T., J. Crouzet, et al. (1990). "Nucleotide sequence, organisation and structural analysis of the products of genes in the *nirB-cysG* region of the *Escherichia coli* K-12 chromosome." Eur J Biochem **191**(2): 315-23.
- Pearson, I. V., M. D. Page, et al. (2003). "A mutant of *Paracoccus denitrificans* with disrupted genes coding for cytochrome c550 and pseudoazurin establishes these two proteins as the in vivo electron donors to cytochrome cd1 nitrite reductase." J Bacteriol **185**(21): 6308-15.
- Pfaltz, A., A. Kobelt, et al. (1987). "Biosynthesis of coenzyme F430 in methanogenic bacteria. Identification of 15,17(3)-seco-F430-17(3)-acid as an intermediate." Eur J Biochem **170**(1-2): 459-67.
- Philippot, L. (2002). "Denitrifying genes in bacterial and Archaeal genomes." Biochim Biophys Acta **1577**(3): 355-76.
- Raux-Deery, E., H. K. Leech, et al. (2005). "Identification and characterization of the terminal enzyme of siroheme biosynthesis from *Arabidopsis thaliana*: a plastid-located sirohydrochlorin ferrochelatase containing a 2Fe-2S center." J Biol Chem **280**(6): 4713-21.
- Raux, E., A. Lanois, et al. (1996). "Salmonella typhimurium cobalamin (vitamin B12) biosynthetic genes: functional studies in *S. typhimurium* and *Escherichia coli*." J Bacteriol **178**(3): 753-67.
- Raux, E., A. Lanois, et al. (1998). "Cobalamin (vitamin B12) biosynthesis: functional characterization of the *Bacillus megaterium* *cbi* genes required to convert uroporphyrinogen III into cobyrinic acid a,c-diamide." Biochem J **335** (Pt 1): 167-73.
- Raux, E., A. Lanois, et al. (1998). "Cobalamin (vitamin B12) biosynthesis: identification and characterization of a *Bacillus megaterium* *cobI* operon." Biochem J **335** (Pt 1): 159-66.
- Raux, E., H. K. Leech, et al. (2003). "Identification and functional analysis of enzymes required for precorrin-2 dehydrogenation and metal ion insertion in the biosynthesis of sirohaem and cobalamin in *Bacillus megaterium*." Biochem J **370**(Pt 2): 505-16.

- Raux, E., T. McVeigh, et al. (1999). "The role of *Saccharomyces cerevisiae* Met1p and Met8p in sirohaem and cobalamin biosynthesis." *Biochem J* **338** (Pt 3): 701-8.
- Raux, E., H. L. Schubert, et al. (1998). "Cobalamin (vitamin B12) biosynthesis--cloning, expression and crystallisation of the *Bacillus megaterium* S-adenosyl-L-methionine-dependent cobalt-precorrin-4 transmethylase CbiF." *Eur J Biochem* **254**(2): 341-6.
- Robin, C., F. Blanche, et al. (1991). "Primary structure, expression in *Escherichia coli*, and properties of S-adenosyl-L-methionine:uroporphyrinogen III methyltransferase from *Bacillus megaterium*." *J Bacteriol* **173**(15): 4893-6.
- Roessner, C. A. (2002). "Use of cobA and cysGA as red fluorescent indicators." *Methods Mol Biol* **183**: 19-30.
- Roessner, C. A. and A. I. Scott (1995). "Fluorescence-based method for selection of recombinant plasmids." *Biotechniques* **19**(5): 760-4.
- Sakakibara, H., K. Takei, et al. (1996). "Isolation and characterization of a cDNA that encodes maize uroporphyrinogen III methyltransferase, an enzyme involved in the synthesis of siroheme, which is prosthetic group of nitrite reductase." *Plant J* **10**(5): 883-92.
- Santander, P. J., C. A. Roessner, et al. (1997). "How corrinoids are synthesized without oxygen: nature's first pathway to vitamin B12." *Chem Biol* **4**(9): 659-66.
- Saunders, N. F., E. N. Houben, et al. (1999). "Transcription regulation of the nir gene cluster encoding nitrite reductase of *Paracoccus denitrificans* involves NNR and NirI, a novel type of membrane protein." *Mol Microbiol* **34**(1): 24-36.
- Sawers, R. G. (1991). "Identification and molecular characterization of a transcriptional regulator from *Pseudomonas aeruginosa* PAO1 exhibiting structural and functional similarity to the FNR protein of *Escherichia coli*." *Mol Microbiol* **5**(6): 1469-81.
- Schon, A., G. Krupp, et al. (1986). "The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA." *Nature* **322**(6076): 281-4.
- Schubert, H. L., E. Raux, et al. (2002). "The structure of *Saccharomyces cerevisiae* Met8p, a bifunctional dehydrogenase and ferrocyclase." *Embo J* **21**(9): 2068-75.
- Schubert, H. L., K. S. Wilson, et al. (1998). "The X-ray structure of a cobalamin biosynthetic enzyme, cobalt-precorrin-4 methyltransferase." *Nat Struct Biol* **5**(7): 585-92.
- Seetharam, B. and D. H. Alpers (1982). "Absorption and transport of cobalamin (vitamin B12)." *Annu Rev Nutr* **2**: 343-69.
- Shemin, D. (1946). "The life span of the red blood cell." *J Biol Chem* **166**: 627-636.
- Shemin, D. (1989). "An illustration of the use of isotopes: the biosynthesis of porphyrins." *Bioessays* **10**(1): 30-5.
- Shima, S., M. Goubeaud, et al. (1997). "Crystallization and preliminary X-ray diffraction studies of methyl-coenzyme M reductase from methanobacterium thermoautotrophicum." *J Biochem* **121**(5): 829-30.
- Silvestrini, M. C., S. Falcinelli, et al. (1994). "*Pseudomonas aeruginosa* nitrite reductase (or cytochrome oxidase): an overview." *Biochimie* **76**(7): 641-54.
- Silvestrini, M. C., M. G. Tordi, et al. (1990). "The reaction of *Pseudomonas* nitrite reductase and nitrite. A stopped-flow and EPR study." *J Biol Chem* **265**(20): 11783-7.
- Smith, A. G. B. and J. A. e. al (1999). "Compartmentation of tetrapyrrole metabolism in higher plants." pp. 281-294.
- Smith, K. M. (1991). Biosynthesis of tetrapyrroles; structure and biosynthesis of bacteriochlorophylls.
- Sofia, H. J., G. Chen, et al. (2001). "Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods." *Nucleic Acids Res* **29**(5): 1097-106.
- Spencer, J. B., N. J. Stolowich, et al. (1993). "The *Escherichia coli* cysG gene encodes the multifunctional protein, siroheme synthase." *FEBS Lett* **335**(1): 57-60.

- Stover, C. K., X. Q. Pham, et al. (2000). "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen." *Nature* **406**(6799): 959-64.
- Stroupe, M. E., H. K. Leech, et al. (2003). "CysG structure reveals tetrapyrrole-binding features and novel regulation of siroheme biosynthesis." *Nat Struct Biol* **10**(12): 1064-73.
- Tanaka, R. and A. Tanaka (2007). "Tetrapyrrole biosynthesis in higher plants." *Annu Rev Plant Biol* **58**: 321-46.
- Timkovich, R., M. S. Cork, et al. (1984). "Proposed structure for the noncovalently associated heme prosthetic group of dissimilatory nitrite reductases. Configuration of substituents of acrylochlorin." *J Biol Chem* **259**(24): 15089-93.
- Timkovich, R., M. S. Cork, et al. (1984a). "Proposed structure for the noncovalently associated heme prosthetic group of dissimilatory nitrite reductases. Identification of substituents." *J Biol Chem* **259**(3): 1577-85.
- Towbin, H., T. Staehelin, et al. (1979). "Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications." *Proc Natl Acad Sci U S A* **76**(9): 4350-4.
- Troup, B., C. Hungerer, et al. (1995). "Cloning and characterization of the *Escherichia coli* hemN gene encoding the oxygen-independent coproporphyrinogen III oxidase." *J Bacteriol* **177**(11): 3326-31.
- Van Spanning, R. J., A. P. De Boer, et al. (1995). "Nitrite and nitric oxide reduction in *Paracoccus denitrificans* is under the control of NNR, a regulatory protein that belongs to the FNR family of transcriptional activators." *FEBS Lett* **360**(2): 151-4.
- Van Spanning, R. J., E. Houben, et al. (1999). "Nitric oxide is a signal for NNR-mediated transcription activation in *Paracoccus denitrificans*." *J Bacteriol* **181**(13): 4129-32.
- Vavilin, D. V. and W. F. Vermaas (2002). "Regulation of the tetrapyrrole biosynthetic pathway leading to heme and chlorophyll in plants and cyanobacteria." *Physiol Plant* **115**(1): 9-24.
- Vevodova, J., R. M. Graham, et al. (2004). "Structure/function studies on a S-adenosyl-L-methionine-dependent uroporphyrinogen III C methyltransferase (SUMT), a key regulatory enzyme of tetrapyrrole biosynthesis." *J Mol Biol* **344**(2): 419-33.
- Vollack, K. U., J. Xie, et al. (1998). "Localization of denitrification genes on the chromosomal map of *Pseudomonas aeruginosa*." *Microbiology* **144** (Pt 2): 441-8.
- Vollack, K. U. and W. G. Zumft (2001). "Nitric oxide signaling and transcriptional control of denitrification genes in *Pseudomonas stutzeri*." *J Bacteriol* **183**(8): 2516-26.
- Waibel, A. E., T. Peter, et al. (1999). "Arctic ozone loss due to denitrification." *Science* **283**(5410): 2064-9.
- Warren, M. J. (2006). "Finding the final pieces of the vitamin B12 biosynthetic jigsaw." *Proc Natl Acad Sci U S A* **103**(13): 4799-800.
- Warren, M. J., E. L. Bolt, et al. (1994). "Gene dissection demonstrates that the *Escherichia coli* *cysG* gene encodes a multifunctional protein." *Biochem J* **302** (Pt 3): 837-44.
- Warren, M. J., E. Raux, et al. (2002). "The biosynthesis of adenosylcobalamin (vitamin B12)." *Nat Prod Rep* **19**(4): 390-412.
- Warren, M. J., C. A. Roessner, et al. (1990). "The *Escherichia coli* *cysG* gene encodes S-adenosylmethionine-dependent uroporphyrinogen III methylase." *Biochem J* **265**(3): 725-9.
- Warren, M. J. and A. I. Scott (1990). "Tetrapyrrole assembly and modification into the ligands of biologically functional cofactors." *Trends Biochem Sci* **15**(12): 486-91.
- Warren, M. J., N. J. Stolowich, et al. (1990). "Enzymatic synthesis of dihydrosirohydrochlorin (precorrin-2) and of a novel pyrrocorphin by uroporphyrinogen III methylase." *FEBS Lett* **261**(1): 76-80.
- Whooley, M. A., McLoughlin, A.J. (1982). "The Regulation of Pyocyanin Production in *Pseudomonas Aeruginosa*." *Eur J Appl Microbiol Biotechnol* **15**: 161 - 166.

- Williams, P. A., V. Fulop, et al. (1997). "Haem-ligand switching during catalysis in crystals of a nitrogen-cycle enzyme." Nature **389**(6649): 406-12.
- Woodcock, S. C. and M. J. Warren (1996). "Evidence for a covalent intermediate in the S-adenosyl-L-methionine-dependent transmethylation reaction catalysed by sirohaem synthase." Biochem J **313** (Pt 2): 415-21.
- Wu, J. Y., L. M. Siegel, et al. (1991). "High-level expression of Escherichia coli NADPH-sulfite reductase: requirement for a cloned cysG plasmid to overcome limiting siroheme cofactor." J Bacteriol **173**(1): 325-33.
- Xiong, J., C. E. Bauer, et al. (2007). "Insight into the haem d1 biosynthesis pathway in heliobacteria through bioinformatics analysis." Microbiology **153**(Pt 10): 3548-62.
- Yamanaka, T. (1992). The Biochemistry of Bacterial Cytochromes. Berlin, Germany, Springer Verlag KG.
- Yap-Bondoc, F., L. L. Bondoc, et al. (1990). "C-methylation occurs during the biosynthesis of heme d1." J Biol Chem **265**(23): 13498-500.
- Ye, R. W., B. A. Averill, et al. (1994). "Denitrification: production and consumption of nitric oxide." Appl Environ Microbiol **60**(4): 1053-8.
- Youn, H. S., Q. Liang, et al. (2004). "Compound 800, a natural product isolated from genetically engineered Pseudomonas: proposed structure, reactivity, and putative relation to heme d1." Biochemistry **43**(33): 10730-8.
- Zajicek, R. S., M. L. Cartron, et al. (2006). "Probing the unusual oxidation/reduction behavior of Paracoccus pantotrophus cytochrome cd1 nitrite reductase by replacing a switchable methionine heme iron ligand with histidine." Biochemistry **45**(37): 11208-16.
- Zumft, W. G. (1993). "The biological role of nitric oxide in bacteria." Arch Microbiol **160**(4): 253-64.
- Zumft, W. G. (1997). "Cell biology and molecular basis of denitrification." Microbiol Mol Biol Rev **61**(4): 533-616.

**A: Summary of phenotypes conferred within bacterial strains or plasmids.**

F'	Host contains an F' episome, with the stated features
<i>dcm</i>	DNA cytosine methylase mutation Blocks methylation of cytosine
$\lambda$ DE3	Bacteriophage lambda carrying the gene for T7 RNA polymerase is integrated into the host genome. The DE3 designation means the strains contain the $\lambda$ DE3 lysogen which carries the gene for T7 RNA polymerase under control of the lacUV5 promoter. IPTG is required to induce expression of the T7 RNA polymerase.
<i>endA1</i>	DNA-specific endonuclease I mutation improves quality of plasmid DNA isolations.
<i>gal</i>	Blocks catabolism of galactose.
<i>gyrA</i> 96	DNA gyrase mutation Confers resistance to nalidixic acid.
<i>hsdR17</i> ( $r_K^-$ , $m_K^+$ )	Host DNA restriction and methylation system mutation: Restriction minus, modification positive or the <i>E. coli</i> K strain methylation system. Allows cloning without cleavage of transformed DNA by endogenous restriction endonucleases. DNA prepared from this strain can be used to transform $r_K^+$ <i>E. coli</i> strains.
<i>hsdS</i> ( $r_B^-$ , $m_B^-$ )	Mutation of specificity determinant for host DNA restriction and methylation system. Restriction minus, modification minus for the <i>E. coli</i> B strain methylation system. Allows cloning without cleavage of transformed DNA by endogenous restriction endonucleases. DNA prepared from this strain is unmethylated by the <i>hsdS</i> 20 methylases
<i>lac</i>	Lactose Promoter
<i>lacIq</i>	Overproduction of the lac repressor protein. Leads to high levels of the lac repressor protein, inhibiting transcription from the lac promoter.

<i>lacZ</i> ΔM15	Partial deletion of β-D-galactosidase gene. Allows complementation of β-galactosidase activity by α-complementation sequence in pGEM Vectors. Allows blue/white selection for recombinant colonies when plated on X-Gal.
LysS	pLysS plasmid is integrated into the host genome. Strains carrying this plasmid will be tet resistant and produce T7 lysozyme, a natural inhibitor of T7 RNA polymerase, thus lowering background transcription of sequences under the control of the T7 RNA polymerase promoter
<i>ompT</i>	Mutation of protease VII, an outer membrane protein. Reduces proteolysis of expressed proteins.
<i>proAB</i>	Mutations in proline metabolism. Requires proline for growth in minimal media.
<i>recA1</i>	Mutation in recombination; minimizes recombination of introduced DNA with host DNA, increasing stability of inserts. Inserts are more stable in recA 1 than recA 13 hosts.
<i>rel A1</i>	ppGpp synthetase I mutation, a novel nucleotide guanosine 5'-diphosphate-3'-diphosphate produced in response to starvation by relA ribosomal protein sensing uncharged tRNA. Allows RNA synthesis in the absence of protein synthesis.
<i>SupE44</i>	Suppressor mutations. Suppresses amber (UAG) mutations.
<i>tac</i>	Tryptophan and Lactose Hybrid Promoter
<i>traD</i> 36	Transfer factor mutation. Prevents transfer of F'episome
Thi	Mutation in thiamine metabolism, Thiamine required for growth in minimal media.

**B: Summary table of properties conferred by the *nir* genes and their encoded proteins:**

GENE	FUNCTION	PRIMARY LOCUS NAME	SWISS-PROT/TrEMBL AC	NCBI ACC NO D84475.1	GENE LENGTH	RANGE ON <i>P. Aeruginosa PA01</i>	GC %	PROTEIN LENGTH	PRT MW Da	PRT MW + Histag 2.182kDa	THEORETICAL PRT PI	PROTEIN EXTINCTION CO-EFF
Nir E	uroporphyrin-III c-methyltransferase-	PA0510	P95417	BAA12682.1	839	573422 to 572586	72.16%	279	29654.11	31.83611	6.6	0.679
Nir J	haem d1 biosynthesis protein	PA0511	P95416	BAA12681.1	1163	574596 to 573436	69.85%	387	43952.7	46.1347	6.9	1.369
Nir H	haem d1 biosynthesis protein	PA0512	P95415	BAA12680.1	515	575105 to 574593	69.20%	171	19581.73	21.76373	9.79	1.51
Nir G	probable transcriptional regulator	PA0513	P95414	BAA12679.	443	575523 to 575083	69.39%	147	16551.86	18.73386	4.71	0.939
Nir L	haem d1 biosynthesis protein	PA0514	P95413	BAA12678.1	524	576040 to 575519	71.07%	174	20210.2	22.3922	9.21	1.335
Nir D	probable transcriptional regulator	PA0515	P95412	BAA12677.1	452	576489 to 576040	70.00%	150	17017.36	19.19936	5.95	0.853
Nir F	haem d1 biosynthesis protein	PA0516	Q51480	AAG03905.	1178	577676 to 576501	69.05%	392	43338.02	45.5202	6.07	1.232

**C: Statement of originality.**

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of others, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline. I acknowledge the helpful guidance and support of my supervisor, Professor Martin Warren.