Siderophore Mediated Iron Acquisition by *Sinorhizobium meliloti*

Thesis presented for the degree of

Doctor of Philosophy

by

Colm Cooke, B.Sc.

under the supervision of

Michael O'Connell, B.A.(Mod), Ph.D.

School of Biotechnology

Dublin City University

Ireland

July 2014

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Degree of Doctor of Philosophy is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____ ID No.: _____

Date: _____

Dedicated to Dr. Michael O’Connell,

Thank you for your advice and friendship these last four years.

Rest in Peace.

Table of Contents

Chapter One

Introduction

1.1 The importance of Iron 2

1.2 Classification of Siderophores 4

1.2.1 Catecholate type siderophores 5

1.2.2 Phenolate type siderophores 5

1.2.3 Carboxylate type siderophores 6

1.2.4 Hydroxamate type siderophore 7

1.2.5 Hydroxamate-citrate siderophores 8

1.2.6 Fungal siderophores 9

1.2.7 NRPS-dependent and NRPS-independent siderophore biosynthesis 11

1.3 Siderophore transport in Gram-negative bacteria 14

1.3.1 TonB-dependent outer membrane receptors 15

1.3.2 The TonB protein complex 19

1.3.3 Plug domain movement and subsequent siderophore transport 20

1.3.4 Siderophore transport across the inner membrane 22

1.4 Iron regulated gene expression 25

1.4.1 Ferric uptake regulator, Fur Regulation 25

1.4.2 Rhizobial iron regulator, RirA regulation 27

1.4.3 Iron responsive regulator, Irr Regulation 28

1.4.4 AraC-like transcriptional regulators 30

1.4.5 Extracytoplasmic function (ECF) sigma factors 31

1.5 Rhizobial-legume symbiosis 32

1.5.1 *Sinorhizobium meliloti* symbiosis 32

1.6 Iron acquisition in free living Rhizobia 35

1.6.1 *Sinorhizobium meliloti* 35

1.6.2 *Rhizobium leguminosarum* 37

1.6.3 *Bradyrhizobium japonicum* 39

1.7 Mass-Spectrometry analysis of siderophores 41

1.8 Summary 43

Chapter Two

Methods and Materials

2.1 Bacterial strains, Primer sequences and Plasmids 47

2.2 Microbiological Media 52

2.3 Solutions and Buffers for DNA analysis 55

2.4 Antibiotics 58

2.5 Isolation and Purification of Nucleic Acids 59

2.5.1 Plasmid DNA isolation using the Genelute Plasmid Miniprep Kit 59

2.5.2 Preparation of total genomic DNA using the Wizard Genomic DNA Kit 59

2.5.3 Purification of DNA using the illustra GFX PCR DNA and Gel Band Purification Kit 60

2.5.4 Tri-reagent RNA extraction 61

2.5.5 cDNA Synthesis from extracted RNA 62

2.6 Quantitation of DNA and RNA 63

2.7 Agarose Gel Electrophoresis 63

2.8 Competent Cells 63

2.8.1 RbCl Competent Cells 64

2.8.2 Preparation of Electrocompetent Cells 64

2.8.3 Transformation of Chemically Competent Cells 65

2.8.4 Transformation of Electrocompetent Cells 65

2.8.5 Determination of Competent Cell Efficiency 65

2.8.6 Sterilisation and Cleaning of Electroportion Cuvettes 66

2.9 Storing of bacterial stocks 66

2.10 Bacterial conjugation by Triparental Mating 66

2.11 Siderophore Detection 67

2.11.1 Chrome Azurol S (CAS) medium (Schwyn and Neilands, 1987) 67

2.11.2 Chrome Azurol S liquid assay 68

2.12 Siderophore production and analysis 69

2.12.1 Production of rhizobactin 1021 69

2.12.2 Purification of rhizobactin 1021 69

2.12.3 Analysis of siderophore by ESI-MS 70

2.13 Enzymatic Reactions 71

2.13.1 Enzymes and Buffers 71

2.13.2 RNase preparation 71

2.13.3 Primer preparation 71

2.13.4 Standard PCR Reaction Mixture 71

2.13.5 Standard PCR Cycling Conditions 72

2.13.6 1 kb Plus DNA Ladder (Invitrogen) 72

2.14 In silico analysis of DNA and protein sequences 73

Chapter Three

The Role of Acetyltransferases in Rhizobactin 1021 Biosynthesis

3.1 Introduction 75

3.2 Comparison of the rhizobactin 1021 biosynthesis operon to similar siderophore operons 80

3.2.1 Comparison of the gene arrangement of biosynthesis clusters 82

3.2.2 Protein homology analysis of the rhizobactin 1021 operon to related siderophore operons 85

3.2.3 Summary of bioinformatic analysis of siderophore biosynthesis proteins 96

3.3 The principles of triparental mating and mutagenesis of *S. meliloti* 97

3.4 Antibiotic resistance cassette mutagenesis of *sma2339* 100

3.4.1 Cloning of *rhtA-sma2339* for the complementation of the S. meliloti 2011*sma2339* mutant 105

3.5 Assessment of rhizobactin 1021 utilisation by the S. meliloti 2011*sma2339* mutant 106

3.6 Assessment of siderophore production by the *S. meliloti* 2011*sma2339* mutant 108

3.7 Antibiotic resistance cassette mutagenesis of *rhbD* 110

3.7.1 Cloning of *rhbDEF* and *rhbEF* for the complementation of the *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* mutants 114

3.8 Assessment of siderophore production by the *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* mutants 115

3.9 Analysis of mutant produced siderophores by ESI-MS 118

3.9.1 Sample preparation and siderophore concentration determination 118

3.9.2 High Pressure Liquid Chromatography of siderophore extracts 120

3.9.3 ESI-MS analysis of siderophore extracts 123

3.10 Summary and Discussion 143

3.11 Conclusions 147

Chapter Four

Coprogen Mediated Iron Acquisition in Sinorhizobium meliloti 2011

4.1 Introduction 153

4.2 Bioinformatic analysis of the pSymA megaplasmid for candidate coprogen receptors 156

4.2.1 Analysis of the genome proximal to *sma1747* 159

4.2.2 Summary of the analysis of the region surrounding *sma1747* 163

4.3 Complementation of S. meliloti Rm818 with *sma1747* and *sma1746-sma1741* 165

4.3.1 Cloning of *sma1747* 165

4.3.2 Cloning of *sma1746-sma1741* 166

4.3.3 Coprogen utilisation by *S. meliloti* Rm818 complemented strains 167

4.4 Antibiotic resistance cassette mutagenesis of sma1747 170

4.4.1 Assessment of *S. meliloti* 2011*rhtX-3sma1747* for coprogen utilisation 174

4.5 Antibiotic resistant cassette mutagenesis of inner membrane coprogen transport systems 175

4.5.1 Note on the mutagenesis of *sma1746* 175

4.5.2 Mutagenesis of *fhuP* in *S. meliloti* Rm818 176

4.5.3 Bioassay assessment of the inner membrane transport systems for coprogen utilisation 178

4.6 Summary and Discussion 179

4.7 Conclusion 181

Chapter Five

Siderophore Mediated Regulation of Outer Membrane Receptor Expression

5.1 Introduction 185

5.2 Bioinformatic analysis of Sma1749 and Smc01610 188

5.2.1 Analysis of Sma1749 188

5.2.2 Analysis of Smc01610 189

5.2.3 Summary of AraC-type protein analysis 189

5.3 Background to coprogen and ferrichrome transport gene expression analysis 190

5.3.1 Background to qPCR 190

5.3.2 Optimisation of qPCR assays 192

5.4 Analysis of the role of coprogen in the expression of *fhuE* and *sma1749* 193

5.4.1 Expression analysis of *fhuE* and *sma1749* 196

5.4.2 Summary of the expression analysis of *fhuE* and *sma1749* 199

5.5 Expression analysis of *fhuA* and *smc01610* 201

5.5.1 Expression analysis of *fhuA* and *smc01610* 102

5.5.2 Summary of the expression analysis of *fhuA* and *smc01610* 207

5.6 Possible modes of action for Sma1749 and Smc01610 209

5.6.1 Analysis of the promoter region of *fhuE* 210

5.6.2 Analysis of the promoter region of *sma1749* 211

5.6.3 Analysis of the promoter region of *fhuA* 211

5.6.4 Analysis of the promoter region of *smc01610* 212

5.6.5 Discussion of the putative binding sites and mode of action 213

5.7 Conclusion 217

Concluding Remarks

6.1 Concluding Remarks 220

References 224

Abstract

Sinorhizobium meliloti 2011 is a free living soil bacterium notable for being the intracellular nitrogen fixing symbiont of *Medicago sativa*. Like most bacteria iron is vital to its survival in both symbiotic and free living conditions. It produces an asymmetrically lipidated siderophore, rhizobactin 1021 (rhz1021), in response to iron limitation. Previous studies have attributed rhz1021 production to the gene products of the operon *rhbABCDEF*. The roles of two acetyltransferases in rhz1021 biosynthesis, *rhbD* and an unstudied gene *sma2339* have been assessed. Analysis shows that *S. meliloti* 2011*sma2339* is capable of producing an intact rhz1021 structure but at a reduced rate to wild type and that the *sma2339* gene can poorly complement *rhbD*. The *rhbD* mutant strain is highly defective in siderophore production confirming its central role in biosynthesis.

In addition to rhz1021 production *S. meliloti* 2011 engages in siderophore pirating as a means of acquiring iron. Analysis of this phenomenon led to the identification of a *fhuE* homologue encoding an outer membrane (OM) coprogen receptor. Coprogen transport across the inner membrane (IM) has proved more complex as redundant transport systems are involved.

S. meliloti 2011 is known to utilise ferrichrome, ferrioxamine B and haem for iron with each solute transported through a dedicated OM receptor. The IM transport system for these three solutes is an integration of elements forming a transport "Split System" which comprises two periplasmic binding proteins, FhuP and HmuT, and an ATPase/transport unit of HmuUV. Data show that all genes involved are up regulated under iron stress with the exception of the ferrichrome OM receptor *fhuA* which was found to be up regulated only in the presence of ferrichrome. The role of an AraC-like gene, *smc01610,* located proximal to *fhuA* in the regulation of ferrichrome uptake has been investigated.

Acknowledgements

This thesis is the culmination of nearly four years work that would not have been possible without the support, friendship and advice of my colleagues, friends and family.

I would foremost like to thank Dr. Michael O'Connell for giving me the opportunity to pursue a Ph.D. as a member of his research group and for his patience, guidance and encouragement for which I will always be grateful. I would also like to thank my two lab colleagues present for most of my time in DCU, Dr. Norah Cassidy for all the good times we had in the lab and post-doctoral researcher Dr. Damien Keogh for his advice and friendship over the years and for his advice at the final stages of my Ph.D. write up.

I would also like to thank all my friends that I have made during my time in DCU, we've had some great fun over the years. In particular in would like to thank Emma for making the last few months the most special and for giving me reasons other than work to come to the lab.

Finally, I thank my parents, Elizabeth and James for giving me all the support I could ever need and to my brothers Seamus, Adrian, Declan and Kieran for all your encouragement throughout the years. This thesis is for you.

List of Figures:

Figure 1.1: Iron chelation moieties associated with siderophore type 4

Figure 1.2: Structure of Enterobactin (Fiedler et al. 2001) 5

Figure 1.3: Structure of yersiniabactin (Fetherston et al. 2010) 6

Figure 1.4: Structure of Staphyloferrin A (Konetschnyrapp et al. 1990) 6

Figure 1.5: Structure of ferrioxamine B and ferrioxamine E (Yamanaka et al. 2005) 7

Figure 1.6: Structures of aerobactin and rhizobactin 1021 (Challis 2005) 9

Figure 1.7: Backbone structure of the ferrichrome siderophores (Renshaw et al. 2002) 10

Figure 1.8: Backbone structures of rhodotorulic acid and coprogen (Renshaw et al. 2002) 10

Figure 1.9: Biosynthesis pathway for ferrichrome A in U. maydis (Winterberg et al. 2010) 11

Figure 1.10: Domain schematic of Fer3, modified from (Winterberg et al. 2010) 12

Figure 1.11: Biosynthesis pathway of aerobactin (Challis 2005) 13

Figure 1.12: Unfolded topology of the FepA protein (Buchanan et al. 1999) 16

Figure 1.13: Ribbon diagram of the structure of the FepA protein (Buchanan et al. 1999) 17

Figure 1.14: Representation of the rhizobactin 1021 regulon 36

Figure 1.15: Representation of haem and xenosiderophore uptake by *S. meliloti* 37

Figure 2.1: Vector map of pBBR1MCS-5 50

Figure 2.2: Vector map of pRK415 50

Figure 2.3: Vector map of pUC4K 51

Figure 2.4: Vector map of pJQ200sk 51

Figure 2.5: 1 kb Plus DNA Ladder 72

Figure 3.1: Structure of rhizobactin 1021 75

Figure 3.2: Organisation of the rhizobactin 1021 biosynthesis and transport genes 76

Figure 3.3: Proposed biosynthesis pathways for rhizobactin 1021 77

Figure 3.4: Core structure of citrate dihydroxamate siderophores 80

Figure 3.5: Biosynthesis operons of citrate dihydroxamate siderophores 83

Figure 3.6: Alignment of the siderophore biosynthesis acetyltransferase proteins 94

Figure 3.7: Alignment of RhbD and Sma2339 protein sequences 95

Figure 3.8: Process of triparental mating 98

Figure 3.9: Schematic of the homologous recombination events 99

Figure 3.10: Schematic for the construction of the construct for mutagenesis of *sma2339* 100

Figure 3.11: Confirmation of the *S. meliloti* 2011*sma2339* mutant by PCR 103

Figure 3.12: Restriction digest confirmation of the *S. meliloti* 2011*sma2339* mutant 104

Figure 3.13: Example of a positive iron nutrition bioassay result 107

Figure 3.14: Analysis of siderophore production by *S. meliloti* 2011*sma2339* 109

Figure 3.15: Schematic of the construction of the construct for mutagenesis of *rhbD* 110

Figure 3.16: Confirmation of the *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* mutants by PCR 113

Figure 3.17: Assessment of siderophore production from *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* 116

Figure 3.18: CAS Assay Standard Curve with Desferrioxamine B 119

Figure 3.19: HPLC Chromatogram of siderophore extracted from *S. meliloti* 2011 120

Figure 3.20: HPLC Chromatogram of siderophore extracted from *S. meliloti* 2011*rhbA62* 121

Figure 3.21: HPLC Chromatogram of siderophore extracted from *S. meliloti* 2011*sma2339* 122

Figure 3.22: Spectrum of the whole extracted siderophore sample from *S. meliloti* 2011 125

Figure 3.23: Spectrum of the isolated rhizobactin 1021 ion fragmentation pattern 127

Figure 3.24: Spectrum of the MS^3 analysis of the rhizobactin 1021 precursor ion 128

Figure 3.25: Spectrum of the whole extracted siderophore sample from *S. meliloti* 2011*sma2339* 130

Figure 3.26: Spectrum of the fragmentation pattern observed from the 538.2 ion in the *S. meliloti* 2011*sma2339* extract 131

Figure 3.27: Spectrum of the isolated 531.3 m/z ion fragmentation pattern from the *S. meliloti* 2011*sma2339* extract 132

Figure 3.28: Spectrum of the isolated 513.1 m/z ion fragmentation pattern from *S. meliloti* 2011*sma2339* extract 133

Figure 3.29: Spectrum of the whole extracted siderophore sample from *S. meliloti* 2011*rhbD* (pCC103) 135

Figure 3.30: Spectrum of the isolated 531.1 m/z ion fragmentation pattern from the *S. meliloti* 2011*rhbD* (pCC103) extract 136

Figure 3.31: Spectrum of the isolated 513.1 m/z ion fragmentation pattern from *S. meliloti* 2011*rhbD* (pCC103) extract 138

Figure 3.32: Spectrum of the whole extracted siderophore sample from *S. meliloti* 2011*rhbA62* 140

Figure 3.33: Spectrum of the isolated 531.3 m/z ion fragmentation pattern from the *S. meliloti* 2011*rhbA62* extract 141

Figure 3.34: Spectrum of the isolated 513.1 m/z ion fragmentation pattern from *S. meliloti* 2011*rhbA62* extract 142

Figure 3.35: Modified rhizobactin 1021 biosynthesis pathway 146

Figure 3.36: Rhizobactin 1021 biosynthesis operon 147

Figure 4.1: Schematic of rhizobactin 1021 and schizokinen uptake 153

Figure 4.2: Schematic of ferrichrome and ferrioxamine B uptake 154

Figure 4.3: Representation of the genome structure of *S. meliloti* 2011 154

Figure 4.4: Structure of coprogen and ferrioxamine B, labelled A and B respectively 156

Figure 4.5: Protein sequence alignment of FoxA, FhuA and Sma1747 158

Figure 4.6: Protein sequence alignment of Sma1747, ShmR and RhtA 158

Figure 4.7: Amino acid sequence of Sma1749 159

Figure 4.8: Amino Acid sequence of Sma1747 160

Figure 4.9: Amino acid sequence of Sma1746 160

Figure 4.10: Amino acid sequence of Sma1745 161

Figure 4.11: Amino acid sequence of Sma1742 161

Figure 4.12: Amino acid sequence of Sma1741 162

Figure 4.13: Amino acid sequence of Sma1740 162

Figure 4.14: Gene arrangement of the possible coprogen uptake system 163

Figure 4.15: Schematic for the creation of the construct used to mutagenize *sma1747* 170

Figure 4.16: Confirmation of the *S. meliloti* 2011*rhtX-3sma1747* mutant by PCR 173

Figure 4.17: Confirmation of the *S. meliloti* Rm818*fhuP* mutant by PCR 177

Figure 5.1: Genome position and gene arrangement of *fhuE, sma1749, fhuA* and *smc01610* 186

Figure 5.2: Protein sequence of Smc01610 189

Figure 5.3: Expression analysis of *fhuE* and *sma1749* in iron deplete conditions 196

Figure 5.4: Expression analysis of *fhuE* and *sma1749* in iron deplete conditions supplemented with 1µM coprogen 197

Figure 5.5: Expression analysis of *fhuE* and *sma1749* in iron replete conditions supplemented with 1 µM coprogen 198

Figure 5.6: Expression analysis of *fhuA* and *smc01610* in iron deplete conditions 202

Figure 5.7: Expression analysis of *fhuA* and *smc01610* in iron replete conditions supplemented with 1 µM ferrichrome 203

Figure 5.8: Expression analysis of *fhuA* and *smc01610* in iron deplete conditions supplemented with 1 µM ferrichrome 204

Figure 5.9: Expression analysis of *fhuA* and *smc01610* in iron deplete conditions supplemented with 100 nM ferrichrome 205

Figure 5.10: Expression analysis of *fhuA* and *smc01610* in iron deplete conditions supplemented with 10 nM ferrichrome 206

Figure 5.11: Schematic of the activity of AraC 209

Figure 5.12: Promter region of *fhuE* 210

Figure 5.13: Promoter region of *sma1749* 211

Figure 5.14: Promoter region of *fhuA* 211

Figure 5.15: Promoter region of *smc01610* 212

Figure 5.16: Schematic of the membrane transport of coprogen and ferrichrome 214

Figure 5.17: Proposed model for *fhuE* and *fhuA* regulation in iron deplete conditions 214

Figure 5.18: Proposed model for *fhuE* and *fhuA* regulation in iron replete conditions in the presence their cognate siderophore 215

Figure 5.19: Proposed model for *fhuE* and *fhuA* regulation in iron deplete conditions in the presence their cognate siderophore 216

List of Tables:

Table 2.1: Bacterial Strains 47

Table 2.2: Plasmids 47

Table 2.3: pJQ200sk Derived Plasmids 47

Table 2.4: pBBR1MCS-5 and pRK415 Derived Plasmids 48

Table 2.5: Primer sequences for cassette mutagenesis 48

Table 2.6: Primer sequences for cloning 49

Table 3.1: Variable features of citrate dihydroxamate siderophores 80

Table 3.2: Siderophore producing strains and the corresponding genome sequence source 81

Table 3.3: Proteins showing significant homology to RhtX 85

Table 3.4: Proteins showing significant homology to RhbA 87

Table 3.5: Proteins showing significant homology to RhbB 87

Table 3.6: Proteins showing significant homology to RhbC 88

Table 3.7: Proteins showing significant homology to RhbF 88

Table 3.8: Proteins showing significant homology to RhbE 89

Table 3.9: Proteins showing significant homology to RhtA 90

Table 3.10: Proteins showing significant homology to RhbD 92

Table 3.11: Proteins showing significant homology to Sma2339 92

Table 3.12: Acetyltranferase proteins BLASTp searched against *S. meliloti* 93

Table 3.13: Zappo colour scheme for multiple alignments 95

Table 3.14: Sequence of the primers used to mutagenize *sma2339* 101

Table 3.15: PCR cycling conditions for the amplification of *sma2339* mutagenesis fragments 101

Table 3.16: Primer sequences for cloning *rhtA-sma2339* into pBBR1MCS-5 105

Table 3.17: PCR conditions for amplification of *rhtA-sma2339* 105

Table 3.18: Analysis of rhizobactin 1021 utilisation by *S. meliloti* 2011*sma2339* 107

Table 3.19: Sequence of the primers used to mutagenize *rhbD* 111

Table 3.20: PCR cycling conditions for amplifing *rhbD* mutagenesis fragments 111

Table 3.21: Primer sequences for the amplification of *rhbDEF* and *rhbEF* for cloning into pBBR1MCS-5 114

Table 3.22: PCR cycling conditions for the amplification of *rhbDEF* and *rhbEF* 114

Table 3.23: Nominal levels of siderophore extracted from rhizobactin mutant strains 119

Table 3.24: Instrument settings used for sample detection and analysis 124

Table 3.25: Explanation of the prominent peaks from the analysis of the *S. meliloti* 2011 126

Table 3.26: Product ions from the fragmentation of the rhizobactin 1021 precursor ion 127

Table 3.27: Product ions from MS^3 of the rhizobactin 1021 precursor ion 129

Table 3.28: Rhizobactin 1021 related m/z values matching values observed in *S. meliloti* 2011 130

Table 3.29: MS^2 product ions from *S. meliloti* 2011*sma2339* matching *S. meliloti* 2011 132

Table 3.30: Observed matches between MS^3 spectra for *S. meliloti* 2011 and

S. meliloti 2011*rhbD* (pCC103) 138

Table 4.1: *S. meliloti* 2011 proteins similar to FoxA as determined by BLASTp 156

Table 4.2: Primer sequences for cloning sma1747 into pRK415 165

Table 4.3: PCR conditions for amplification of *sma1747* 166

Table 4.4: Primer sequences for the cloning of *sma1746-sma1741* into pBBR1MCS-5 166

Table 4.5: PCR conditions for the amplification of *sma1746-sma1741* 167

Table 4.6: Iron nutrition bioassay analysis of complemented *S. meliloti* Rm818 167

Table 4.7: Sequence of the primers used to mutagenize *sma1747* 171

Table 4.8: PCR cycling conditions for amplifying *sma1747* mutagenesis fragments 171

Table 4.9: Assessment of coprogen utilisation by *S. meliloti* 2011*rhtX-3sma1747* 174

Table 4.10: Primer sequences for confirmation of *S. meliloti* Rm818*fhuP* 176

Table 4.11: PCR cycling conditions for confirmation of *S. meliloti* Rm818*fhuP* 176

Table 4.12: Assessment of coprogen utilisation by *S. meliloti* Rm818*fhuP* 178

Table 4.13: List of inner membrane transport associated genes up regulated in iron deplete conditions 179

Table 4.14: *S. meliloti* 2011 proteins similar to FhuE from *E. coli* K12 by BLASTp 181

Table 5.1: List of primers used in to monitor the gene expression of *fhuA*, *fhuE* and associated genes 192

Table 5.2: Expected product size and primer concentrations used for each qPCR assay 192

Table 5.3: Growth conditions for expression analysis of *fhuE* and *sma1749* 193

Table 5.4: Growth conditions for expression analysis of *fhuA* and *smc01610* 201

Abbreviations, Units and Prefixes

Abbreviations

ABC	ATP-dependent Binding Cassette
ATP	Adenosine Triphosphate
CAS	Chrome Azurol Sulphonate
CCRH	Colonised Curled Root Hair
Ct	Cycle Threshold
DNA	Deoxyribonucleic Acid
ECF	Extra Cytoplasmic Function
EDTA	Ethylenediaminetetraacetic acid
EPS	exopolysacharride
ESI-MS	Electrospray Ionisation Mass Spectrometry
FAB-MS	Fast Atom Bombardment Mass Spectrometry
HPLC	High-Pressure Liquid Chromatography
d/sH_2O	Deionised/Sterile H_2O
ICE	Iron Control Element
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
IRO	Iron-Responsive Operator
LB	Lysogeny Broth
MCS	Multiple Cloning Site
MFS	Major Facilitator Superfamily
NMR	Nuclear Magnetic Resonance
dNTP	Deoxyribonucleotide triphosphate

NRPS	Non-Ribosomal Peptide Synthesis
OD	Optical Density
PBS	Phosphate Buffered Saline
PBP	Periplasmic Binding Protein
PCR	Polymerase Chain Reaction
qPCR	Quantitative PCR
pI	Isoelectric Point
RT	Reverse Transcription
SB	Super Broth
TAE	Tris Acetate EDTA
TBDT	TonB-Dependent Transporter
TCA	Tricarboxylic Acid Cycle
TE	Tris-EDTA
TY	Tryptone Yeast
Amp^R	Ampicillin Resistant
Cm^R	Chloramphenicol Resistant
Gm^R	Gentamicin Resistant
Km^R	Kanamycin Resistant
Sm^R	Streptomycin Resistant
Tc^R	Tetracycline Resistant

Units

Å	Angstrom
Da	Dalton

°C	Degrees Celsius
g	Gram
V	Volt
hr	Hour
kb	Kilobase pair
L	Litre
m/z	Mass to Charge Ratio
Mb	Megabase pair
min	Minute
M	Molar
psi	Pounds per Square Inch
pH	Power of Hydrogen
sec	Second
rpm	Revolutions per Minute

Prefixes

M	Mega (10^{6})
k	Kilo (10^{3})
m	Milli (10^{-3})
µ	Micro (10^{-6})
n	Nano (10^{-9})
p	pico (10^{-12})

Publications

Papers

Cooke C., Keogh D., O Cuív P., O'Connor B., Kelleher B. and O'Connell M. Investigation of the role of acetyltransferases RhbD and Sma2339 in the biosynthesis of rhizobactin 1021 from *S. meliloti* 2011, Molecular Microbiology, in preparation.

Cooke C., Keogh D., O'Connor B. and O'Connell M. Identification of FhuE, the outer membrane receptor for coprogen in *S. meliloti* 2011 and siderophore mediated regulation of *fhuE* expression, Journal of Bacteriology, in preparation.

Selected Posters

Investigation of the lipid acylation of the siderophore rhizobactin 1021, 1st EMBO | EMBL Symposium: New Approaches and Concepts in Microbiology, Heidelberg, Germany, October 2013.

Characterisation of novel transport systems for the utilisation of hydroxamate type siderophores by *Sinorhizobium meliloti* 2011 and *Pseudomonas aeruginosa* PAO1, 8th International Biometals Symposium (Biometals 2012), Brussels, Belgium, July 2012

Expression analysis of a xenosiderophore acquisition system in *Sinorhizobium meliloti* 2011, Young Microbiologists Symposium, UCC, Cork Ireland, June 2012

RT-qPCR analysis of the expression of a novel transporter system for the acquisition of haem, ferrichrome and ferioxamine B in *Sinorhizobium meliloti*, EMBO Meeting, Vienna Austria, Sept 2011

Chapter One

Introduction

1.1 The importance of Iron

The two lipid membranes possessed by Gram-negative bacteria provide an excellent protective barrier against the external environment. Bacteria possessing an outer membrane are more resilient to harsh environmental conditions due to the outer membrane insulating the cytoplasmic membrane. However, the presence of this outer membrane does introduce complications especially in the area of nutrient and mineral transport as there are two barriers for solutes to traverse. Most solutes are transported through the outer membrane by porins that allow movement through the membrane into the periplasmic space followed by internalisation into the cell by dedicated inner membrane transport systems. Expression of these transporters and associated proteins is driven by the cellular need for the cognate solute and therefore only transporters that are required are present. The activity of transporters, transcription machinery and other associated cellular processes are inexorably coupled to cellular metabolic rate. The primary mineral element associated with cellular processes is iron as the ease with which it participates in redox reactions makes it a favourable electron donor/acceptor in metabolic processes.

Bacteria mainly contain iron in the redox centre of redox enzymes which allows the generation of voltage potentials ranging from -300 mV to +700 mV (Andrews, Robinson and Rodriguez-Quinones 2003). This ability to contribute to the formation of large redox potentials facilitates the activity of a vast number of proteins and enzymes that require an electrical potential to function. The primary examples of where such enzymes are found are in DNA synthesis, cellular metabolism, oxygen transport via haem molecules, photosynthesis, and nitrogen fixation and these examples demonstrate how crucial iron is to cellular survival. There is a significant limitation to this reliance on iron as it is extremely insoluble in oxygenated and mild pH conditions. This results in soluble iron levels of only 10^{-9} M to non-existent which is far below the 10^{-7} M required to support microbial growth (Braun and Hantke, 2013). Iron scarcity is due to ferrous iron, Fe^{2+} becoming oxidised to the insoluble ferric form, Fe^{3+} in the presence of oxygen and compounded by H_2O resulting in the formation of $Fe(OH)_3$.

The insolubility of iron poses a significant obstacle to the survival of microorganisms and has led to them developing dedicated iron acquisition systems, the most prominent being siderophore production. Siderophores are low molecular weight, <2000 Da

secondary metabolites with an extremely high affinity for Fe^{3+} (Budzikiewicz, 2010). They are secreted by microorganisms in response to iron stress, chelate iron from the environment and are subsequently internalised. High affinity dedicated transport systems identify the siderophore-ligand complex and internalise it into the cell. Arising from their efficiency and the requirement of most organisms for iron, siderophore transport systems are widely found across microorganisms. Study of these systems has led to the identification of the responsible transport proteins many of which have been subsequently characterised at a genetic and biochemical level. These dedicated systems in Gram-negative bacteria generally comprise a high affinity outer membrane receptor and a less specific periplasmic binding protein dependent ABC type transport system (Schalk and Guillon 2013). However, recently there have been numerous examples of Major Facilitator Superfamily (MFS) transporters facilitating the internalisation of siderophores across the inner membrane indicating that inner membrane siderophore transport is not as functionally singular as outer membrane transport (Funahashi et al. 2013; Chatfield et al. 2012; O Cuiv et al. 2007; O Cuiv et al. 2004).

The transport of siderophores has been well characterised in many organisms and is essential for these organisms to thrive. Equally important to understanding how microorganisms obtain iron is investigation of siderophore biosynthesis. There are two methods employed by microorganisms to assemble siderophores, the non-ribosomal peptide synthetase (NRPS) method and the NRPS-independent biosynthesis method. The NRPS-dependent method relies on large multi-domain biosynthesis enzymes that catalyse the assembly of the constitutive compounds of the siderophore by transferring the siderophore intermediate from one step to the next in a modular fashion. This allows siderophores to be built up stepwise resembling a production line and upon completion of the backbone structure the siderophore is released and modified through the activity of associated tailoring enzymes. The NRPS-independent method relies on a suite of biosynthesis proteins that first modify precursor molecules. This is followed by recognition by siderophore synthetases which catalyse their assembly into the mature siderophore structure (Barry and Challis 2009; Challis 2005; Crosa and Walsh 2002).

1.2 Classification of Siderophores

Siderophores are a structurally divergent class of secondary metabolites linked via their very high affinities for Fe^{3+}. This structural diversity results from there being a variety of producers each with its own requirements to thrive in distinct environmental niches. Siderophores are produced by terrestrial soil bacteria, plant symbionts, animal and plant pathogens, commensals, marine bacteria and numerous strains of fungi each of which tailors siderophores differently to suit the producer's needs. It is estimated that over 500 different siderophores exist with 270 having been structurally characterised. Such a number of metabolites carrying out a similar function has led to a nomenclature based solely on the chemical moieties involved in iron binding as this is the feature of siderophores showing the greatest conservation (Hider and Kong 2010).

Classification based on iron binding moieties has allowed all siderophores to be described via four distinct types namely, catecholate type, phenolate type, hydroxamate type and the carboxylate type with a number of mixed types also in existence. Figure 1.1 is a representation of the chemical groups associated with each siderophore type.

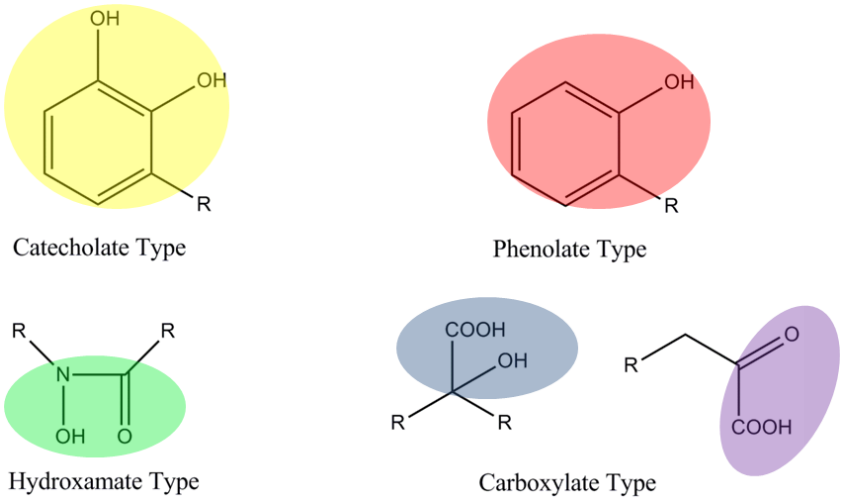


Figure 1.1: Iron chelation moieties associated with siderophore type

Each of the four types of siderophore has had members of its group chemically characterised. Prominent siderophores of each type will be discussed here namely enterobactin for the catcholate type, yersiniabactin for the phenolate type, ferrioxamine B for the hydroxamate type and staphyloferrin A for the carboxylate type. Mention will also be given to relevant mixed type siderophores and fungal siderophores (Miethke and Marahiel 2007).

1.2.1 Catecholate type siderophores

Enterobactin, also known as enterochelin is the characteristic siderophore produced by the family Enterobacteriaceae. It was first identified as a product from *S. typhimurium* where it was termed enterobactin and subsequently identified in *E. coli* extracts where it was referred to as enterochelin (Pollack and Neilands 1970; O'Brien, Cox and Gibson 1970). Assembly of the final enterobactin structure is achieved in a NRPS-dependent manner resulting in the formation of an amide linkage between 2,3-dihydroxybenzoic acid with L-serine which then is formed into a trimer. Under ideal binding conditions enterobactin has the highest known affinity for ferric iron however this is undermined by the hydrolysis of the trimeric structure at slightly acidic pH (Raymond, Dertz and Kim 2003; Neilands 1981). The structure of enterobactin is given in figure 1.2.

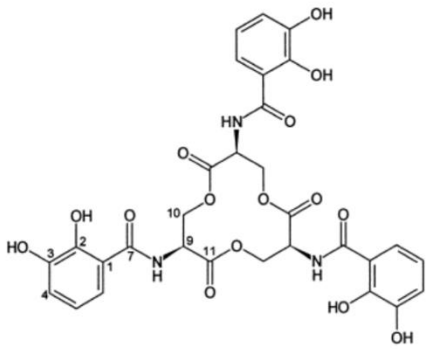


Figure 1.2: Structure of Enterobactin (Fiedler et al. 2001)

Enerobactin is an example of the cyclic catecholate type siderophores, agrobactin produced by the plant pathogen *Agrobacterium tumefaciens* is an example of a linear catecholate type siderophore. It was characterised as a threonyl peptide of spermidine with three residues of 2,3-dihydroxybenzoic acid and an oxazoline ring (Ong, Peterson and Neilands 1979).

1.2.2 Phenolate type siderophores

Yersiniabactin is the primary siderophore produced by *Yersinia pestis*, the causative agent of the bubonic and pneumonic plagues and is produced under iron stress (Perry et al. 1999). It is synthesised in a NRPS-dependent manner and comprises one molecule of salicaylate, three molecules of cysteine and a malonyl-CoA which are sequentially assembled into the final structure (Pfeifer et al. 2003). It has been found to be a

virulence factor for *Y. pestis* and knockout of the biosynthesis pathway attenuates the pathogen (Fetherston et al. 2010). The structure of yersiniabactin is given in figure 1.3.

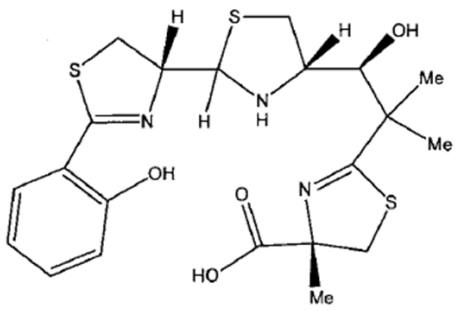


Figure 1.3: Structure of yersiniabactin (Fetherston et al. 2010)

1.2.3 Carboxylate type siderophores

Staphyloferrin A was first identified as produced by *Staphylococcus hyicus* DSM 20459 when grown under iron deficient conditions (Konetschnyrapp et al. 1990). The biosynthetic genes were subsequently identified in *S. aureus* and mapped to the *sfa* gene cluster which encodes for an NRPS-independent biosynthesis cluster (Beasley et al. 2009). A biosynthesis pathway has yet to be proposed but the original structural analysis found that ornithine enhanced production indicating that it may be a constituent of the siderophore. Figure 1.4 gives the structure of staphyloferrin A.

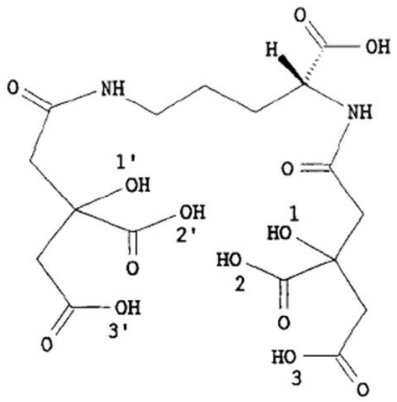


Figure 1.4: Structure of Staphyloferrin A (Konetschnyrapp et al. 1990)

1.2.4 Hydroxamate type siderophore

Hydroxamate siderophores are found to be synthesised by both microbial and fungal strains. They contain multiple bidendate iron coordinating groups that act together to chelate iron.

The ferrioxamines are a diverse group of siderophores in themselves but share a core structure and differ in side groups. They possess a low molecular weight and have a trihydroxamate based structure. They have been shown to be produced by numerous strains most notably *Streptomyces coelicolor*, *Streptomyces pilosus* and *Erwinia amylovora* along with various other *Streptomyces* strains (Patel, Song and Challis 2010; Kachadourian et al. 1996; Feistner, Stahl and Gabrik 1993; Muller and Raymond 1984,). A model for the biosynthesis of the linear compound ferrioxamine B and cyclical compound ferrioxamine E has been proposed by (Barona-Gomez et al. 2006). The two pathways overlap extensively with one another in that the first step is the conversion of L-Lysine to N-hydroxy-1,5-diaminopentane by the activity of the DesA and DesB proteins. This compound is then differentially acetylated by the DesC protein with either an acetyl-CoA or a succinyl-CoA moiety with the produced compounds acting as the primary substrates for the siderophores. The assembly of the siderophore is completed by the DesD protein with the resultant siderophore structure dependent on which of the DesC produced molecules is incorporated. The structures of the two siderophores ferrioxamine B and ferrioxamine E are given in figure 1.5.

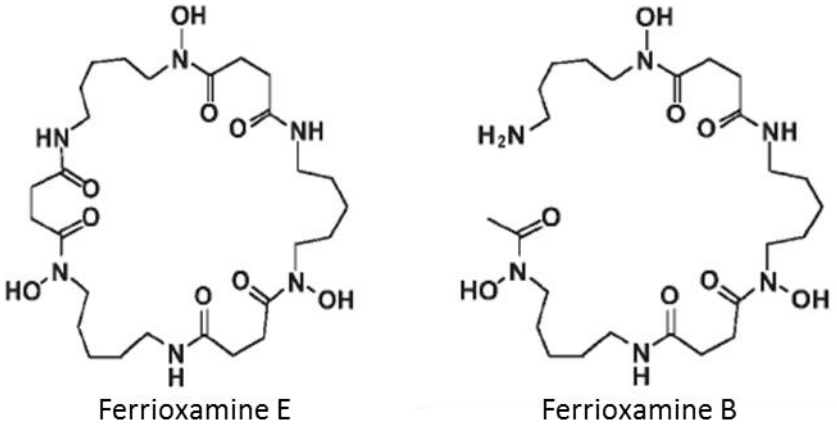


Figure 1.5: Structure of ferrioxamine B and ferrioxamine E (Yamanaka et al. 2005)

The cyclic structure of ferrioxamine E and the linear structure of ferrioxamine B are evident from figure 1.5.

The ferrioxamine siderophores are noteworthy outside of the area of microbial iron acquisition as the methane-sulfonate derivative of ferrioxamine B marketed under the name Desferral has been used for the treatment of hemochromatosis in humans (Flaten et al. 2012). It is administered via a subcutaneous injection as it is susceptible to acidic degradation. It chelates excess iron in the blood and is subsequently removed from the body by the kidneys.

1.2.5 Hydroxamate-citrate siderophores

The only mixed type siderophores that will be described here are the hydroxamate-citrate siderophores. These are based on a citrate backbone with the final structure being created through the substitution of the terminal citrate carboxyl groups with hydroxamate groups resulting in a higher affinity for iron. Only a brief introduction to this class of siderophore is presented here as an in depth analysis of this siderophore type is carried out in Chapter 3.

The first discovered citrate-hydroxamate siderophore was isolated from iron depleted supernatants of *Enterobacter aerogenes* by (Gibson F. 1969). The biosynthesis of aerobactin was elucidated by work carried out on *E. coli* harbouring the pColV-K30 plasmid and it was found to be assembled in an NRPS-independent manner (de Lorenzo et al. 1986; de Lorenzo and Neilands 1986). The two main substrates for aerobactin biosynthesis are a modified L-Lysine structure and a citrate molecule.

There are numerous other hydroxamate-citrate based siderophores in existence many of which will be discussed in Chapter 3. However due to the importance of rhizobactin 1021 to the subsequent work presented in this thesis it is important for it to be introduced. Rhizobactin 1021 is the sole siderophore produced by *Sinorhizobium meliloti* 1021. Similar to aerobactin it has a backbone of citrate however in place of L-Lysine derived hydroxamate structures there are 1,3-diaminopropane derived structures (Persmark et al. 1993). The other key difference in the structures of the two siderophores is that rhizobactin 1021 is adorned asymmetrically with a decenoic acid moiety at the terminal of one of the 1,3-diaminopropane molecules where aerobactin has only a basic acetyl group. The biosynthesis genes of rhizobactin 1021 are striking similar to that for aerobactin with the main deviation being the presence of two genes

encoding for 1,3-diaminopropane production in the rhizobactin 1021 operon (Lynch et al. 2001). The structures of both these siderophores are given in figure 1.6.

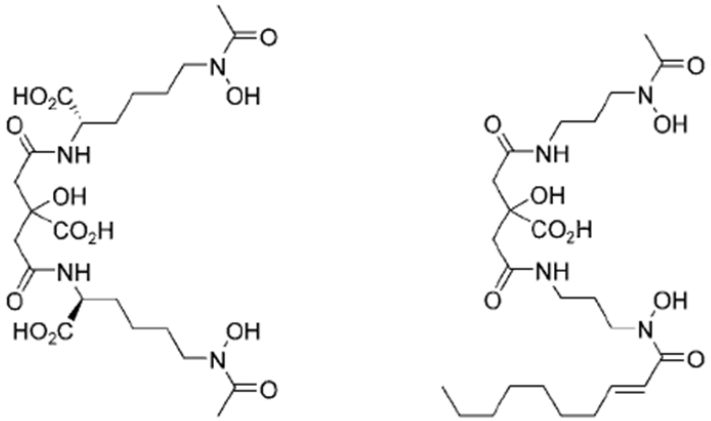


Figure 1.6: Structures of aerobactin and rhizobactin 1021 (Challis 2005)

1.2.6 Fungal siderophores

There are a large number of fungal siderophores that display a diverse range of structures. Most prominent amongst these siderophores are the ferrichromes, coprogens, fusarinines and polycarboxylates and the first three will be discussed here in further detail (Winkelmann 2007).

The fungal siderophores are of great interest not only due to their structural diversity but also due to the quantities at which they are present in the environment. Two studies into the natural levels of siderophore in the soil have indicated that fungal siderophores of the ferrichrome and ferricrocin families can be found in concentrations up to 10 nM (Essen et al. 2006; Prabhu, Biolchini and Boyer 1996). Ferrichrome was first isolated from supernatant extracts from *Ustilago sphaerogena* with a few other producers subsequently identified (Emery 1971). The biosynthesis of ferrichrome has also being described in *U. sphaerogena* and was found to be assembled in a NRPS-dependent manner (Winterberg et al. 2010; Yuan et al. 2001; Mei, Budde and Leong 1993). Biosynthesis appears to be limited to fungal species however a large number of bacteria can utilise it through the Fhu or related transport system. Ferrichromes display variation that is due to the precursor molecules being derived from various amino acids namely glycine, serine and alanine. Variation is also introduced through the hydroxamate

residue structures having functional groups substituted with carboxylic acid groups. The backbone structure of the ferrichromes is given in figure 1.7.

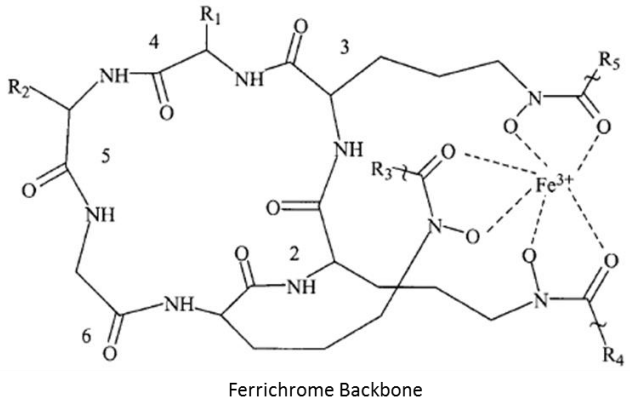


Figure 1.7: Backbone structure of the ferrichrome siderophores (Renshaw et al. 2002)

The R groups in figure 1.7 denote variable regions in the structure with each variation constituting a different ferrichrome.

The fusarinines are found as monomers, linear dimers or trimers, or cyclic trimers. Each monomer segment consists of a hydroxyornithine molecule acetylated with an anhydro-mevalonic acid residue (Winkelmann 1990). The monomers of the *cis* and *trans* fusarinine are the structural units of a number of fungal siderophores. The coprogen like siderophores are linear dihydroxamate and trihydroxamate ligands that are composed of *trans*-fusarinine units (Leong and Winkelmann 1998). Depending on the arrangement of the fusarinine units when present as a dimer the resulting siderophores are either rhodotorulic acid or dimerum acid. Figure 1.8 gives the structures of the backbone of rhodotorulic acid and also of coprogen.

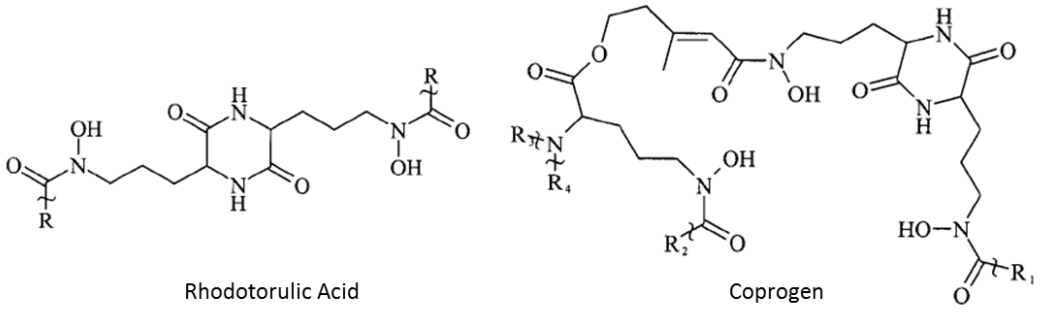


Figure 1.8: Backbone structures of rhodotorulic acid and coprogen (Renshaw et al. 2002)

The R in figure 1.8 can represent a number of different functional groups from basic methyl groups to entire amino acid derived structures.

1.2.7 NRPS-dependent and NRPS-independent siderophore biosynthesis

The two methods by which siderophores are synthesised are described as either being NRPS-dependent or NRPS-independent. To highlight the differences and similarities between the two biosynthesis processes a characterised example of each type will be described in detail. The biosynthesis of ferrichrome from *U. maydis* is a prime example of NRPS-dependent biosynthesis and aerobactin biosynthesis is a clear example of NRPS-independent biosynthesis.

As mentioned earlier the biosynthesis of ferrichrome was elucidated by (Winterberg et al. 2010; Yuan et al. 2001; Mei, Budde and Leong 1993). *U. maydis* produces two siderophores, ferrichrome and ferrichrome A. The biosynthesis of ferrichrome A will be described here as it has been completely characterised where only the initiator and NRPS protein for ferrichrome biosynthesis has been identified. The biosynthesis pathway of ferrichrome A is outlined in figure 1.9.

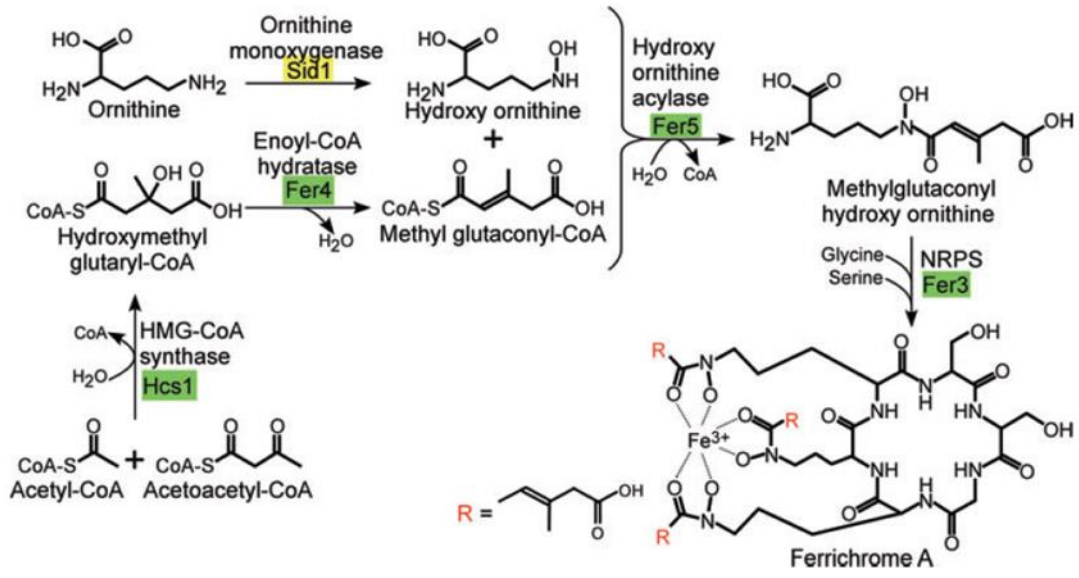


Figure 1.9: Biosynthesis pathway for ferrichrome A in *U. maydis* (Winterberg et al. 2010)

The initiator of biosynthesis is the Sid1 protein as it acts to oxygenate ornithine to hydroxy-ornithine. This step is also the initiator step in ferrichrome biosynthesis. The proteins Hcs1 and Fer4 act in succession to assemble and modify acetyl-CoA and

acetoacetyl-CoA into a molecule of methyl-glutaconyl-CoA. This large CoA associated molecule is fused onto the hydroxy-ornithine by the acetyltransferase activity of Fer5 resulting in the creation of methylglutaconal-hydroxy-ornithine. This final ornithine derivative is assembled into the mature ferrichrome A structure through the activity of the NRPS Fer3. The ferrichrome structure is built up in a stepwise manner with the sequential addition of one glycine and two serine residues which form the cyclical internal structure of ferrichrome A. To understand the process by which the NRPS Fer3 achieves this, the authors undertook an in-depth analysis of the domain structure of Fer3. Figure 1.10 is a schematic of the Fer 3 protein.

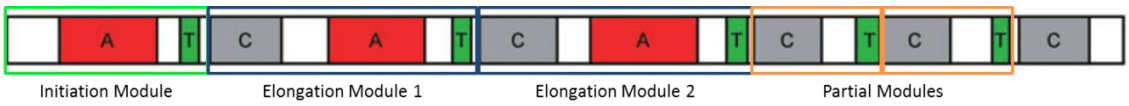
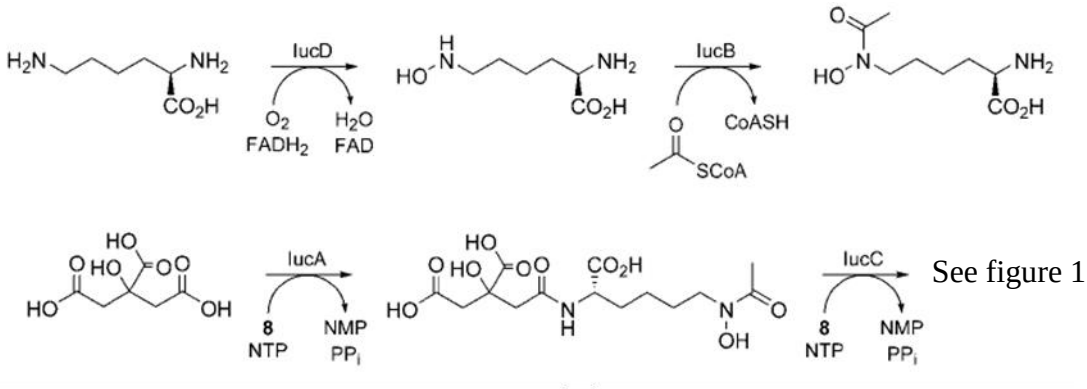


Figure 1.10: Domain schematic of Fer3, modified from (Winterberg et al. 2010)

The Fer3 protein is a very large protein with an amino acid length of 4830. It possesses a number of repeating domains which are labelled A, T and C in figure 1.10. The amino acid to be added to the siderophore is first adenylated by an A domain followed by transference of this activated amino acid to a peptidyl carrier protein by reactive thioester formation carried out by a T domain. This peptide is then bonded to the substrate held at the C domain by the condensation activity of the C domain. These domains are split into a number of modules with the initiation module which modifies the first amino acid boxed in green in figure 1.10. This is then added to the core structure by elongation modules, boxed in blue in figure 1.10, which also prepares the next amino acid and the process is repeated in the next elongation module. There are two partial elongation modules at the C-terminus of the protein, marked in orange which are thought to add the third amino acid with the assistance of one of the adenylation domains of a previous module. As is evident the NRPS proteins are a complex multi-domain protein that acts in a modular fashion to build the siderophore from the cognate precursors.

The biosynthesis of aerobactin has been determined by de Lorenzo et al. (1986) and de Lorenzo and Neilands (1986). It consists of four proteins acting in sequence to produce precursor molecules which are then assembled by the activity of synthetase enzymes. The biosynthesis pathway is given in figure 1.11.



See figure 1.6

Figure 1.11: Biosynthesis pathway of aerobactin (Challis 2005)

The first step in the production of aerobactin is the modification of L-lysine by the oxygenase-like activity of IucD. This structure is then acetylated by the activity of an acetyltransferase, IucB to form N_6-acetyl-N_6-hydroxylysine. This along with citrate is then recognised by IucA to form a citrate-hydroxamate intermediate. This citrate-hydroxamate intermediate is recognised by IucC which adds a second N_6-acetyl-N_6-hydroxylysine to the citrate forming the citrate-dihydroxamate siderophore aerobactin (structure given in figure 1.6).

The key differences between the biosynthesis pathways lie in the final method of assembly. The initial steps involve the preparation of precursor molecules by acetyltransferase-like and oxygenase-like functions. Depending on the complexity of the precursor molecule there may be more preparation steps. The assembly of the siderophore in the case of the NRPS-dependent pathway is achieved by the activity of a large multifunctional protein, whereas assembly of NRPS-independent siderophores is achieved by the activity of synthetase proteins that are relatively small, ~600 amino acids in comparison to ~4500 amino acids for the NRPS proteins. It is unknown whether the synthetases act sequentially to one another or in unison as a protein complex. Perhaps NRPS-independent siderophores represent a more economical method of producing a siderophore as they do not require the translation and processing of such large proteins.

1.3 Siderophore transport in Gram-negative bacteria

Gram-negative bacteria are encapsulated by two lipid bilayers referred to as the outer and inner (cytoplasmic) membranes. The intervening region between these membranes is known as the periplasm and is composed of a number of features that allow the cell to protect itself from possible toxins and facilitate transport of required nutrients. A thin layer of peptidoglycan is also present in the periplasm that gives the cell rigidity as well as providing additional protection. The presence of two protective membranes allows the cell greater control over the substrates that are internalised. Only small uncharged molecules such as O_2, CO_2 and H_2O can pass through the membrane by simple osmosis. Small charged soluble molecules such as Na^+ and Ca^{2+} are transported over the outer membrane by simple porins that allow the transport of molecules based on charge but restrict large molecules from entering (Nikaido 2003). However recent studies in *Bradyrhizobium japonicum* have shown that MN^{2+} transport across the outer membrane is facilitated by a specific transporter MnoP which demonstrates that solute transport can be achieved in a more directed manner (Hohle et al. 2011). Larger molecules are transported over the outer membrane by dedicated substrate specific transporters which recognise and facilitate the transport of molecules such as sucrose and maltose. These porins are intricately structured to only allow exact structures to pass through the pore thereby preventing the accidental uptake of large solutes into the periplasm. These two types of porins function as facilitators as they do not require energy to transport molecules from the environment into the periplasm however to transport non-uniform molecules such as siderophores energy is required. Siderophore transport over the outer membrane is achieved by TonB-dependent outer membrane receptors also known as ligand gated receptors. The energy for these transporters must be transduced across the periplasm from the cytoplasm as there is no energy producing mechanism present in the periplasm. This is achieved through the TonB protein complex (Koebnik, Locher and Van Gelder 2000). Once a substrate is present in the periplasm it is recognised by an inner membrane transport system that completes the internalisation of the substrate.

The area of solute transport in Gram-negative bacteria is too vast to be covered in total in this review and as a result only mechanisms involved in the uptake of siderophores will be described. This will be discussed in terms of the most well understood examples of Gram-negative siderophore utilisation with special mention of unique or unusual deviations from the standard models.

1.3.1 TonB-dependent outer membrane receptors

The TonB-dependent outer membrane receptors or TonB-dependent transporters (TBDT) are a distinctive class of outer membrane transporters. They possess a large diameter pore that is held closed by an N-terminus plug domain. Without this plug domain the pore would be too large to discriminate between substrates that should and should not be transported. As mentioned previously this family of transporters is dependent on energy transduced from the inner membrane through the TonB protein complex. This energy is required to introduce a conformational change in the plug region that allows uptake of the cognate substrate to occur.

The mode of action of the TBDT has been studied in depth based on crystal structures obtained from a variety of their members. Numerous TBDT involved in siderophore uptake have been crystallised and structurally characterised. Amongst these there are examples of receptors from various organisms with specificities for different siderophores namely; FauA from *Bordetella pertussis*, FepA, FecA and FhuA from *E. coli,* and FpvA and FptA from *P. aeruginosa*. Many of these proteins have been crystallised in different stages of interaction with their cognate siderophore and with the TonB protein. The cognate siderophores for the proteins named above are alcaligin for FauA, enterobactin for FepA, citrate for FecA, ferrichrome for FhuA, pyoverdin for FpvA and pyochelin for FptA (Brillet et al. 2009; Cobessi et al. 2005; Cobessi, Celia and Pattus 2005; Cobessi, Celia and Pattus 2004; Yue, Grizot and Buchanan 2003; Ferguson et al. 2002; Buchanan et al. 1999; Ferguson et al. 1998; Locher et al. 1998,). Each of these structures was found to rely on the same basic themes both in their structure and in their function. The pore of the porin comprises 22 anti-parallel β-sheets that arrange themselves in the outer membrane in a barrel arrangement therefore leading to the term β-barrel. Remarkably the β-sheets appear to be tilted at ~45 degrees in each structure. This β-barrel forms a large elliptical pore that generally has dimensions of between 40-46 Å and 24-45 Å for each cross section. To prevent this large pore from acting in an indiscriminate manner it is occluded from the periplasmic N-terminus face by a large plug structure that is positioned in the lumen of the β-barrel. This plug consists of numerous alpha helices and β-sheets that arrange themselves into an ordered structure preventing diffusion through the pore. In addition to the β-barrel and plug it was found that the TonB energising protein interacts with a disordered N-terminal

region located near or in the cytoplasm which is referred to as the TonB box. Figure 1.12 is a representation of an unfolded FepA protein from *E. coli.*

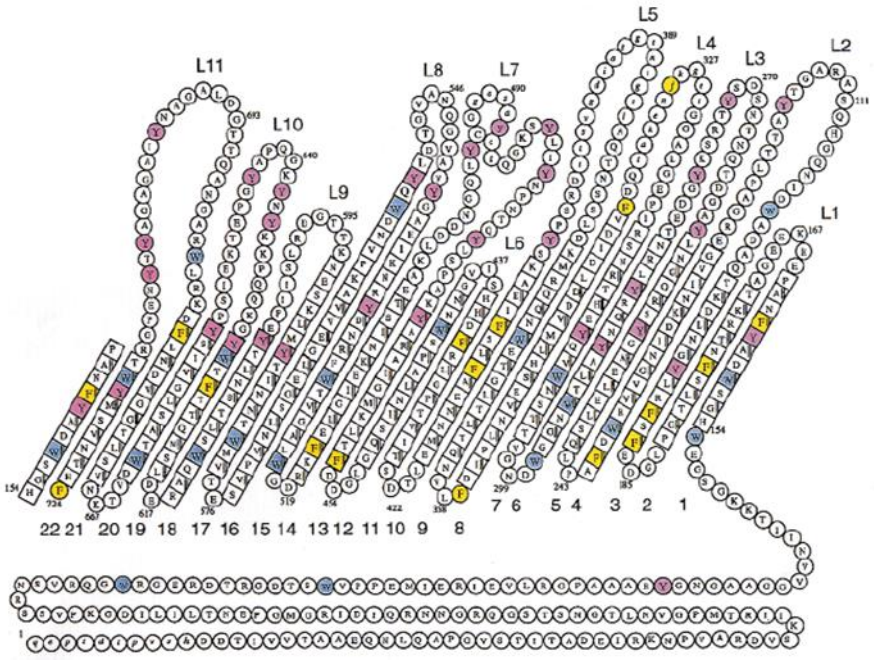


Figure 1.12: Unfolded topology of the FepA protein, (Buchanan et al. 1999)

Figure 1.12 serves to clarify the description of the β-barrel structure of the TBDT proteins. The squares represent the β-sheet forming residues with the circles indicating the rest of the amino acids. The orientation of the protein in the membrane is represented in the diagram with the labelled loops L1-L11 being presented to the external environment. The plug domain structure is not represented in this diagram but would comprise the region between amino acid 1 and amino acid 154. A striking feature of the structure is the large loops that extend out from the protein and these loops play an essential part in the capturing and subsequent identification of the cognate siderophore. To aid in the description of the process of substrate identification a ribbon diagram of the FepA protein is given in figure 1.13.

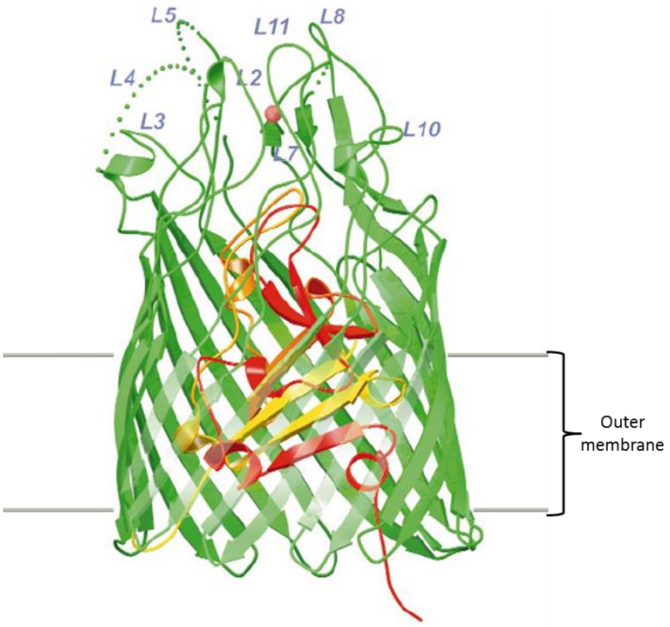


Figure 1.13: Ribbon diagram of the structure of the FepA protein, (Buchanan et al. 1999)

The ribbon diagram in figure 1.13 shows how the structure given in figure 1.12 assembles itself in the membrane. The plug domain, shown as yellow and red in figure 1.13, is positioned so that it interacts with the extracellular environment and it is this along with the presence of the large loop structures that provide the points of interaction between the siderophore and the receptor. The crystal structure of FepA with enterobactin bound was inconclusive with regard to the precise residues involved in siderophore binding however they were able to conclude that it occurred through the loop structures. Crystal structures obtained by Locher et al. (1998) of the FhuA protein in *E. coli* allowed the locations of the siderophore recognition residues to be determined. They obtained crystals of the protein in both the ferrichrome bound and ferrichrome free states. The key residues are located on the apical turns of the plug domain and on the two loops L3 and L11. The siderophore is held through hydrogen bond formation between itself and the receptors binding residues which forms a binding pocket into which ferrichrome is a perfect fit. Other studies carried out on the FhuA protein from *E. coli* determined the binding residues for the siderophore phenylferricrcin and the sideromycin albomycin (Ferguson et al. 2000). Correlating with the observations made in the previous study these two ligands are found in the described binding pocket with the ligands being held by hydrogen bond formation or van der Waals interactions. They identified three residues from the apices of the plug namely Arg81, Gln100 and Tyr116 to be involved in the binding pocket along with

numerous residues from the loop structures. Analysis of citrate binding by the FecA protein found a similar situation in which siderophore binding was achieved by residues on the apices of the plug and on two of the external loops. Additional binding sites for citrate were identified on the β-barrel located near the plug apices (Yue, Grizot and Buchanan 2003). Studies into siderophore binding by the *P. aeruginosa* receptors FptA and FpvA also indicate the involvement of the plug, loops and residues from the internal β-barrel structure (Greenwald et al. 2009; Cobessi et al. 2005; Cobessi, Celia and Pattus 2005; Cobessi, Celia and Pattus 2004). It is clear that recruitment of the siderophore by TBDT proteins is achieved by multiple interactions between the siderophore and the receptor. Upon binding of the substrate a conformational shift occurs that facilitates the interaction with TonB eventually resulting in siderophore transport.

The conformational shift that allows TonB to activate the receptor has been observed in a number of crystal structures. The TonB box is a short conserved sequence that is the main point of interaction between TonB and the receptor and is located at the N-terminal of the receptor. According to crystallography this can be either tucked up into the β-barrel or be located in the periplasm, however without a siderophore being bound to the receptor it is not available to TonB. The conformational changes induced by substrate binding have been observed in the FecA protein and the ShuA haem/haemoglobin receptor from *Shigella dysenteriae* (Cobessi, Meksem and Brillet 2010; Yue, Grizot and Buchanan 2003). It was observed that the loops not involved in binding fold over the top of the receptor which results in binding of the ligand by a number of new binding sites. This contributes to the unfolding of the TonB box, also referred to as disordering which leaves the TonB protein free to interact and transduce energy from the inner membrane.

1.3.2 The TonB protein complex

The TonB protein functions as part of an inner membrane protein complex that also includes ExbB and ExbD referred to as the TonB complex. The sole purpose of the TonB complex is to transduce energy created at the inner membrane by proton motive force to the outer membrane. This must be done as the periplasm has no means of generating a potential gradient and does not contain ATP. The ExbB protein is 26 kDa and traverses the inner membrane via three alpha helices with the majority of the soluble fraction being contained in the cytoplasm. The ExbD protein is 17 kDa and traverses the membrane via one alpha helix and has a short C-terminal present in the periplasm. It is anchored in the cytoplasm by two short alpha helices and a 5-stranded β-sheet. The TonB protein itself is 26 kDa and is comprised of three functional domains. The N-terminal domain acts to anchor the protein in the inner membrane through an alpha helix structure and also contains a signal sequence that directs the protein to the inner membrane. The second domain comprises residues 66-149 and is located in the periplasm. This is a proline rich region and is made up of a series of proline and glutamic acid residues followed by several proline-lysine repeats. The C-terminal of the second domain comprising residues 103-149 gives a degree of flexibility to the TonB protein that facilitates interactions with the outer membrane proteins. The third domain consisting of residues 150-239 form the structure responsible for interacting with its target proteins in the outer membrane (Krewulak and Vogel 2011).

The mode of action of the TonB protein is subject to much debate. Previous models pertaining to how it interacted with the TBDT while spanning the periplasm suggested that it forms a dimer at the C-terminal of the protein and also disengages from the inner membrane entirely. Both of these theories have being disproven as re-interpretation of old results along with the accumulation of new results has provided support for different models. The previous model for TonB was that for the protein to reach the outer membrane it must disengage from the inner membrane and shuttle across the periplasm to interact with the TBDT proteins. This was supported by the fact that the length of the TonB protein, ~100 Å is significantly shorter than the gap between the inner and outer membranes which is estimated at ~180 Å. Further support for the shuttling model was provided by the fact that a truncated TonB missing the proline spacer was still functional. However, experiments in which TonB was forcibly maintained in the inner membrane by fusing a ToxR protein fused to the cytoplasmic N-terminus showed no

deleterious effect on functionality. This along with the discovery that the experimental design of the experiments showing shuttling by protein location were flawed, proves that TonB is associated with the inner membrane at all times. The flaw in the original experiment was identified by the authors a number of years after publication as being the dye Oregon Green 488 maleimide which was thought to be incapable of diffusion through membranes but in fact could. The second reassignment of TonB activity was the discovery that it does not form dimers *in vivo*. It was originally thought to form dimers based on crystal structures of TBDT proteins that had a dimeric TonB C-terminal attached. However *in vivo* crosslinking experiments showed no evidence of dimer structures showing that only mutated TonB proteins form dimers *in vivo* (Gresock et al. 2011; Postle et al. 2010).

These findings have opened up a plethora of possible modes of action for TonB. As the protein is not long enough to span the periplasm when the C-terminus is structured it is now believed that the C-terminus becomes disordered to interact with the TBDT proteins. This model, proposed by (Gresock et al. 2011) and is based on results obtained (Chimento, Kadner and Wiener 2005) suggests that only modest force is required to dislodge the plug domain from a TBDT protein. This was concluded based on comparative structural analysis of four TBDT proteins which found that the interface between the plug domain and the β-barrel is extensively hydrated. This would facilitate large conformational shifts with minimal external force required. This along with the positioning of the TonB box would allow the plug to be dislodged by a mechanism similar to a lever in which a small force translates into a large movement.

1.3.3 Plug domain movement and subsequent siderophore transport

The final step in transporting a solute through a TBDT protein is the displacement of the plug and movement of the solute into the periplasmic space. The mechanics behind this displacement are the subject of much debate. There are two schools of thought; the first indicates that the plug is ejected from the β-barrel which allows the substrate to translocate through the porin and the second suggesting that the pore merely moves to one side slightly allowing the substrate to slide past it.

It is established that the plug is loosely packed into the lumen of the β-barrel structure and that relatively minor force is required to cause a conformational shift (Chimento, Kadner and Wiener 2005). It has been postulated from some crystal structures that a basic conformational shift in which the plug moves to one side forming a relatively small pore is sufficient to allow the solute passes through. This model is also supported by the fact that a small shift can occur at less metabolic cost to the cell than a large conformational shift (Buchanan et al. 1999; Ferguson et al. 1998; Locher et al. 1998). However, this model does not account for how such large molecules such as colicins which utilise TBDT proteins to gain access to cells can fit. Colicins, which can have a molecular weight up to 69 kDa have been found to be transported through FhuA (Wiener et al. 1997). Such a large protein could not be accommodated by such a modest conformational change in the plug region which gives rise to the possibility of the plug being completely ejected into the periplasm in order to complete transport.

Many attempts to elucidate the conformational shift that occurs in the plug during transport have being made. Experiments using the TBDT proteins FhuA and FepA where double cysteine mutants were created to tether the plug to the β-barrel found that if the tether was close to the N terminus of the plug domain transport did not occur but was only reduced when tethered elsewhere (Endriss et al. 2003). Further to this data, labelling experiments based on the accessibility of cysteine residues introduced into the lumen of the β-barrel by mutagenesis found that cysteine residues introduced near the location of the plug in FepA were poorly labelled when no substrate was present however upon addition of enterobactin the labelling increased while transport was occurring. These results support the hypothesis that the plug exits the lumen either partially or fully during transport. The presence of the plug prevents labelling of the cysteine residues when no enterobactin is present and when enterobactin is present labelling increases showing that the plug is no longer present to block labelling (Devanathan and Postle 2007; Ma et al. 2007).

Further work is required to prove the model for the plug being ejected from the lumen during transport however the accumulated results appear to support this model. Regardless of the exact shift in the plug domain the process culminates in the delivery of the cognate solute to the periplasm. It is then sequestered by an inner membrane transporter that completes the internalisation into the cell.

1.3.4 Siderophore transport across the inner membrane

Transport of siderophores across the inner membrane is a much less specific process than transport across the outer membrane. Siderophores primarily utilise periplasmic binding protein (PBP) dependent ABC transporters to cross the inner membrane. These systems generally consist of a periplasmic binding protein that sequesters the siderophore in the periplasm and in turn delivers it to the inner membrane complex which comprises a permease component and an ATPase component. In addition to the ABC transporter dependent method numerous examples of single unit transporters belonging to the (MFS) have been found to transport siderophores across the inner membrane. This method of siderophore transport is less well described at a structural level than the ABC transporter method but uncovers interesting possibilities for siderophore uptake.

1.3.4.1 Siderophore uptake via ABC transporters

The ABC transporter system usually comprises five functional domains; the periplasmic binding protein, two transmembrane proteins that act together to form the permease and two nucleotide binding proteins that act to hydrolyse ATP for energy. Generally the permease unit comprises a homodimer but it can be formed by a heterodimer with the ATPase dimer generally being formed by a homodimer. In *E. coli* the enterobactin ABC transporter has a stereochemistry of $FepBC_2D_2$ with FepB being the PBP, FepC being the permease monomer and FepD being the ATPase monomer which is also the arrangement found for enterobactin ABC transporter in *Vibrio anguillarum*, the $FatBC_2D_2$ complex (Shea and Mcintosh 1991). The permease is usually formed via dimer formation of the two permease subunits however investigation of the hydroxamate siderophore transport system in *E. coli* FhuDBC found that the permease FhuB fused into a single polypeptide chain that functions in the same manner as the usual dimer formation. It was found that the PBP FhuD and the ATPase FhuC function as expected in that FhuD acts independently and FhuC forms a dimer to function (Mademidis and Koster 1998; Mademidis et al. 1997). Further variation in the permease assembly is found in the vibriobactin and enterobactin ABC transporters in *Vibrio cholera* known as VctPDGC and ViuPDGC respectively. Similar to other systems the PBP acts as a monomer and the ATPase functions as a homodimer. The permease in both of these systems in composed of the VctDG and ViuDG proteins for vibriobactin

and enterobactin respectively and they assemble as a heterodimer to function (Wyckoff and Payne 2011; Wyckoff, Mey and Payne 2007). In addition to variations in the permease assembly and structure an example of shared ATPases between siderophore uptake systems has also being identified. The FatBCDE and FvtBCDE systems of *Vibrio anguillarum* pJM1 function to transport anguibactin, and vanchrobactin/enterobactin respectively. It was found that mutation of FatE did not abolish anguibactin uptake while a double mutant of FatE and FvtE did abolish anguibactin transport. This showed that in the absence of the FatE ATPase the FvtE ATPase was capable of interacting with the rest of the anguibactin transport system resulting in restoration of function (Naka, Liu and Crosa 2013).

Comparison of the ATPase components of the ABC transporters showed a number of conserved features; the Walker A motif (GxxGxGKS/T where x is any amino acid), the Walker B motif (hhhD where h is any hydrophobic amino acid), a signature sequence found to be uniquely associated with ABC transporter ATPase proteins (LSGGQQ/R/K/KQR) and the Q-loop that has a conserved Gln residue (Krewulak and Vogel 2008). However no crystal structure of an ABC system ATPase has been solved as yet. The crystal structure of the FhuD PBP from *E. coli* has been solved and the structure was found to be pincer like in that it comprises two lobes that are separated by a cleft that harbours the substrate binding site (Clarke et al. 2002; Clarke et al. 2000).

As mentioned previously inner membrane transport is not as specific as the outer membrane transport by TBDT proteins. Generally a TBDT protein will only transport one type of siderophore along with closely related derivatives of that siderophore. A number of phages and toxins also utilise them but these are not their intended targets as the TBDT protein is hijacked in these circumstances. This is exemplified by the FhuDBC system in *E. coli* where the four TBDT proteins FhuA, FhuE, IutA and complemented FoxA utilise FhuDBC to internalise the siderophores ferrichrome, coprogen, aerobactin and ferrioxamine B respectively (Braun, Hantke and Koster 1998). A second example of ABC systems transporting numerous substrates is the split system for haem and hydroxamate siderophore uptake in *S. meliloti*. It was found that the HmuUV inner membrane permease and ATPase could transport haem via the HmuT PBP and ferrichrome and ferrioxamine B via the FhuP PBP (O Cuiv et al. 2008). In contrast to the highly diverse hydroxamate system in *E. coli, Vibrio cholerae* displays redundancy at the inner membrane for the siderophores vibrobactin and enterobactin

both of which can be transported via the VctPDGC or ViuPDGC ABC systems (Mey et al. 2002).

1.3.4.2 Siderophore uptake via single unit tranporters

There are numerous examples of single unit transporters of the MFS family transporting siderophores across the inner membrane. The first member of this family to be identified was RhtX, the inner membrane transporter for rhizobactin 1021 in *S. meliloti* (O Cuiv et al. 2004). The FptX protein which functions in pyochelin transport in *P. aeruginosa* was also discovered as part of this work and was confirmed by Michel, Bachelard and Reimmann (2007). A second single unit transporter was identified in *P. aeruginosa* named FoxB which is of interest as it does not fit the MFS transporter model and may be an independent class of transporter. This protein was identified through complementation experiments carried out in a *S. meliloti rhtX* mutant with FoxB being able to complement the function of RhtX in *S. meliloti* along with conferring the ability to transport the hydroxamate siderophores ferrichrome and ferrioxamine B to *S. meliloti* mutants deficient in their transport across the inner membrane. An allelic exchange mutant in of FoxB *P. aeruginosa* showed no detectable phenotype suggesting a high degree of redundancy for these siderophores (O Cuiv et al. 2007). In addition to *S. meliloti* and *P. aeruginosa,* siderophore transporting MFS proteins have being identified for acinetoferrin uptake by *Acinetobacter haemolyticus* and for legiobactin uptake by *Legionella pneumophila* (Funahashi et al. 2013; Chatfield et al. 2012). These two proteins were identified as LbtC for legiobactin uptake and ActC for acinetoferrin uptake.

As the number of MFS proteins discovered to be involved in siderophore transport continues to grow it may be that they are not as rare as initially thought. Although the rate of identification has increased there is still very little known about the mode of action of these proteins. They appear to function independently of a PBP component which raises the question of how the siderophore docks with the transporter. Also there is little known on how the siderophore translocates through the permease as no crystal structures have being solved. This family of transporters shares some similarity with the AmpG from *E. coli* that is involved in muropeptide transport for cell wall recycling (Jacobs et al. 1994). This protein relies on proton motive force for energy and it would

be of great interest to investigate the reliance of the siderophore transporters on proton motive force.

1.4 Iron regulated gene expression

There are numerous regulators involved in the control of iron uptake and homeostasis in Gram-negative bacteria. The most prevalent amongst these is the Fur protein which was first identified in *E. coli*. As iron regulation was studied in more depth across a greater number of species other distinct regulators were identified as the regulator of iron homeostasis such as RirA and Irr. These regulators act in a global sense in that they monitor the general iron levels in the cell and act accordingly to cope with low iron levels and high iron levels. In addition to global regulation there are a number of regulatory elements that act in a localised manner and these are represented by AraC like proteins and two component associated sigma factors. These local regulators generally act to enhance expression of key elements involved in iron acquisition.

1.4.1 Ferric uptake regulator, Fur Regulation

The Fur protein is the defining protein of the superfamily of transcriptional regulators known as the FUR superfamily. Members of this family are structurally related and are usually involved in metal homeostasis regulation and by extension oxidative stress regulation as metal ions exacerbate free radical formation. The members of this family that are known to regulate metal homoeostasis are Fur for iron, Zur for zinc, Nur for nickel, Mur for manganese, and Irr for iron but senses through the iron status of cellular haem rather than iron levels directly (Fillat 2014).

The Fur protein has been extensively studied as it is the major iron homeostasis protein in the model organism *E. coli*. It is known to exert control over a broad range of genes many of which are involved in iron uptake, siderophore production and iron storage but surprisingly it also influenced genes related to energy production, pathogenicity and redox-stress response. It controls this regulon by acting as an iron dependent transcriptional repressor which results in gene repression in the presence of high iron concentrations and de-repression when cellular iron levels are deficient (McHugh et al. 2003a). A complete mechanism behind its activity has still to be uncovered however

there is a substantial amount of evidence that allows a relatively complete model to be created. It acts primarily by sensing the levels of iron by directly binding Fe^{2+} (Bagg and Neilands 1987). Upon binding of iron the protein recognises a conserved sequence present upstream or overlapping with the target genes promoter known as the Fur Box or Iron Box. Fur forms a dimer on this Fur Box which results in the repression of transcription. The Fur Box is conserved amongst many organisms that rely on Fur for iron regulation and in *E. coli* it has been found to comprise of the 19 bp inverted repeat consensus sequence GATAATGATAATCATTATC (Escolar, Perez-Martin and de Lorenzo 1998). In addition to this primary function of repression of gene expression through DNA binding it has been found that Fur also functions at a post transcriptional level through the small RNA RhyB. The RhyB sRNA functions through the RNA chaperone Hfq to downregulate genes that are positively regulated by Fur such as iron storage proteins, a superoxide dismutase and genes involved in the TCA cycle, all of which require iron or store iron. Fur prevents RhyB from functioning by binding its promoter and repressing the expression of the sRNA. This results in the translation of iron storage proteins, oxidative stress proteins and proteins with an iron requirement, all of which are beneficial in a high iron situation. It also results in the immediate degradation of iron storage mRNA and the mRNA of proteins with a high iron demand in times of iron starvation (Masse and Gottesman 2002).

Structural studies of Fur have yielded valuable insights into the domains involved in its activity. The N-terminus of the protein folds into a helix-turn-helix structure that allows the protein to bind and recognise the DNA binding motif of the Fur Box (Hantke 2002). The C-terminus of the protein is responsible for iron binding and dimerization which is achieved through the presence of a histidine cluster and four cysteine residues. Mutagenesis of these cysteine residues found that residues at position C_{92} and C_{95} were responsible for the binding of Zinc which enables correct folding of the protein while the residues at position C_{132} and C_{137} were non-essential (Lee and Helmann 2007). The crystal structure of Fur from *P. aeruginosa* has being determined providing important insights into the protein. The DNA binding domain is comprised of four α-helices and two anti-parallel β-sheets with only one of the helices active in DNA binding. The dimerization domain is structured as three anti-parallel β-sheets that occlude a single α-helix. There are two metal binding sites on the *P. aeruginosa* Fur protein known as site one and site two, with the first site found to be capable of binding Zn^{2+} or Fe^{2+} and the

second only binding Zn^{2+} (Lee and Helmann 2007; Pohl et al. 2003). This has led to the first binding site being designated as the iron sensing region.

Fur homologs are found across a number of bacterial species including both Gram-negative and Gram-positive species. Whereas many of these examples are involved in iron homeostasis there is crossover between many proteins in the FUR superfamily. This is the observed case for the Rhizobia as the most prominent homolog to Fur was found to actually be acting like a Mur protein in that it regulated manganese homeostasis (Fillat 2014). This led to the identification of the RirA protein, a member of the Rrf2 family which was found to fulfil the role of Fur in Rhizobial species (Johnston et al. 2007).

1.4.2 Rhizobial iron regulator, RirA regulation

The RirA protein was first identified in *Rhizobium leguminosarum* and subsequently identified in *S. meliloti* as the main global regulator of iron homeostasis (Chao et al. 2005; Viguier et al. 2005; Wexler et al. 2003; Todd et al. 2002). There is little sequence homology between RirA and Fur as RirA is a member of the Rrf2 protein family. The Rrf2 family of regulators has been previously characterised as playing a role in regulating genes encoding for a cytochrome in *Desulfovivrio*. Two other members of this family NsrR and IscR have being shown to regulate nitrogen oxide metabolism in *Nitrosomonas* and regulation of FeS cluster formation respectively (Johnston et al. 2007).

The homologue to Fur in *R. leguminosarum* and *S. meliloti* was found to actually be Mur and acted in response to cellular manganese levels not iron (Diaz-Mireles et al. 2005; Chao et al. 2004; Platero et al. 2004). This regulator was found to control the *sitABCD* operon which encodes for a metal uptake system most likely for manganese.

The RirA protein does not recognise the Fur Box, instead it acts through regions in the promoters of RirA-repressed genes known as iron-responsive operators (IRO). The identification of these motifs provides insights into how RirA represses genes in iron replete conditions (Yeoman et al. 2004). No IRO motif has been identified in *S. meliloti* which leaves the mode of action of RirA in *S. meliloti* unclear. In addition to the global repression of iron associated genes RirA is capable of effecting change on a more local

level. RirA represses the rhizobactin 1021 regulon under iron replete conditions. Included in this regulon is an AraC-like regulator RhrA which is also repressed by RirA. When the cell encounters iron deplete conditions the promoters for rhizobactin 1021 biosynthesis and *rhrA* are de-repressed. Once activated the RhrA protein acts as an activator of the rhizobactin 1021 biosynthesis operon along with the outer membrane receptor *rhtA* (Viguier et al. 2005). This is an example of how iron acquisition can be both globally and locally regulated resulting in a significant response to iron deprivation.

1.4.3 Iron responsive regulator, Irr Regulation

As mentioned earlier Irr is classified as a member of the FUR superfamily of regulators. Homologs have being identified in *B. abortus, A. tumefaciens, R. leguminosarum* and *B. japonicum* (Fillat 2014). The most well characterised Irr protein is from the rhizobial species *B. japonicum.* Extensive research into the role of Irr has found that it plays a very similar role in *B. japonicum* as Fur does in *E. coli* however the mode of action is very different. As described above Fur in *E. coli* acts through sensing the internal status of iron in the cell by the presence of Fe^{2+} whereas Irr senses through the oxidation status of haem in the cell.

Irr is functional under conditions of iron limitation and has been found to act as both a positive and negative regulator of gene expression. Under iron limitation it is found to activate the expression of a number of iron acquisition genes and repress numerous genes involved in iron dependent processes such as the TCA cycle. It also prevents haem biosynthesis under iron limited conditions (Yang et al. 2006; Hamza et al. 1998). Irr recognises and binds a partially conserved binding motif displaying the consensus sequence, TTTAGAANNNTTCTAAA, which is referred to as the iron control element (ICE). Irr binding to this recognition sequence is sufficient to repress gene expression but no molecular basis for activation has been proposed as yet (Sangwan, Small and O'Brian 2008; Rudolph et al. 2006).

Regulation of Irr in *B. japonicum* is primarily achieved through selective protein stability. This stability is regulated in a haem-dependent manner that results in the degradation of Irr in iron replete conditions. This is achieved by direct binding of Irr to the two redox forms of haem the ferrous and ferric forms and through the activity of the

ferrichelatase HemH involved in the final step of haem biosynthesis, the conversion of protoporphyrin IX to haem. Irr binds ferric haem through a haem regulatory motif found near the N-terminus with ferrous haem binding occurring at a histidine pocket located close to the C-terminus. Binding of the two forms of haem allows Irr to sense the iron status of the cell and when iron is plentiful, represented by a greater abundance of ferrous haem, the HemH protein acts to degrade Irr (Yang et al. 2006; Yang, Ishimori and O'Brian 2005; Qi and O'Brian 2002; Hamza and O'Brian 1999). Irr controls haem biosynthesis by repression of the *hemB* gene that encodes for a δ-Aminolevulinic acid dehydratase that is the second step in haem biosynthesis thus ensuring that it is not degraded in a haem-dependent manner under iron limited conditions. It also switches off haem biosynthesis as the deferrated haem precursor is toxic and it is detrimental to have this present in high levels when iron is not available. Perhaps it is not coincidental that the HemH protein converts the haem precursor to haem as well as functioning in the degradation of Irr in iron replete conditions (Hamza et al. 2000). A role for haem in the control of Irr from *R. leguminosarum* has also being identified however it is limited to interfering with activity rather than degradation (Singleton et al. 2010).

The role of Fur in *B. japonicum* has also been investigated and has yielded another variation on the role of this protein in iron and indeed metallo-homeostasis in general. Fur in *B. japonicum* has been found to be iron responsive but it regulates a set of genes that are not related to iron homeostasis instead regulating genes involved in CO_2 fixation and carbon metabolism (Yang, Sangwan and O'Brian 2006). This is a very different role than that observed in other rhizobia where Fur homologs are actually Mur proteins. In addition to this global regulation of genes in response to iron the *B. japonicum* Fur protein acts to regulate Irr at the transcriptional level. Fur binds the Irr promoter in iron replete conditions resulting in a 3 fold reduction in *irr* expression (Hamza et al. 2000). But the role of Fur in *B. japonicum* is more expansive than just iron mediated regulation as it also responds to cellular manganese levels which results in a dual regulation of iron uptake in *B. japonicum* mediated through manganese and iron cellular levels. Under high iron, high manganese conditions Fur binds manganese and binds to the *irr* promoter region therefore repressing transcription. When conditions switch to low iron, high manganese the Irr protein occupies the *irr* promoter therefore preventing the manganese-Fur complex from binding which results in *irr* expression. When iron is high and manganese low, Irr is degraded in the normal haem-dependent

manner and Fur does not bind the *irr* promoter as it is not activated by manganese (Hohle and O'Brian 2010).

1.4.4 AraC-like transcriptional regulators

The AraC protein is a transcriptional regulator that controls the catabolism of arabinose in *E. coli*. It functions as a dimer that is capable of selective binding of DNA binding motifs located proximal to promoter regions of its target genes. Upon sensing the presence of arabinose in the cell the AraC dimer can switch binding sites and activate the expression of the *araBAD* genes. AraC has also being found to repress the *araBAD* genes in the absence of arabinose. This family and related families of transcriptional activators achieve this through direct binding of the substrate, arabinose in the case of AraC, which is carried out by the N-terminus of the protein. Once the substrate has been bound the protein then binds its cognate DNA binding sites through the helix turn helix structure formed by the C-terminus (Schleif 2010; Gallegos et al. 1997).

Many species of bacteria possess AraC-like transcriptional regulators and there are a number of examples where these play a role in regulating genes involved in iron acquisition. One such example is the PchR protein found in *P. aeruginosa* which regulates the outer membrane receptor for pyochelin uptake, FptA. Mutagenesis of *pchA* resulted in the reduction of *fptA* expression in a strain capable of producing pyochelin. Analysis of *fptA* expression from a pyochelin deficient derivative of the background for the *pchA* knockout showed nearly no *fptA* expression. A double mutant resulting in the knockout of *pchA* and pyochelin production restored *fptA* expression to the levels observed in the single *pchA* mutation. These results indicated that in the absence of PchA the *fptA* gene was expressed at low basal levels. In the presence of PchA without pyochelin the PchA protein repressed *fptA* to nearly undetectable levels showing that PchR is a repressor in the absence of the siderophore. In the presence of PchR and pyochein *fptA* expression was strongly activated indicating that PchR acts as an activator in the presence of its target siderophore pyochelin (Heinrichs and Poole 1996). Also the authors reported a Fur Box close to the *fptA* promoter which would suggest that basal level expression is attributed to Fur in the absence of PchA repression.

There are quite a large number of AraC-like transcriptional regulators positioned proximal to genes associated with iron uptake. This would suggest that AraC-like

control of siderophore uptake is not uncommon and would represent an efficient and tightly controlled method of acquiring iron that would enhance the activity of the global iron response regulators previously discussed.

1.4.5 Extracytoplasmic function (ECF) sigma factors

Numerous examples of ECF sigma factor regulation of iron uptake have been described. Perhaps the most extensively studied example of this method of regulation is the control of FecA by the FecIR two component system in *E. coli*. The FecIR system comprises a sigma factor, FecI and an inner membrane associated periplasmic signalling protein FecR. This system acts as a sensory mechanism that responds to ferric citrate to allow the expression of the TBDT protein FecA. FecA is the cognate receptor for ferric citrate and functions in the same manner as other TBDT proteins. Upon binding of ferric citrate a conformational change is transduced through the molecule which signals for the binding of TonB to energise the system. However, the FecR protein spans the periplasm and also recognizes this conformational change by interacting with the N-terminus of FecA. Upon recognition, FecR in turn undergoes a structural change that results in the proteolytic degradation of a small cytoplasmic region of FecR by the site-2 protease RseP. The cleaved peptide binds to FecI resulting in activation. This activated FecI protein binds to the transcriptional start site of *fecA* and recruits the RNA polymerase thus affecting transcription (Braun and Mahren 2005). Phenomena closely related to this mechanism also occur in *P. aeruginosa* and *Bordetella bronchiseptica* (Draper et al. 2011; Vanderpool and Armstrong 2004).

1.5 Rhizobial-legume symbiosis

The process of entering and correctly achieving a functional symbiosis is dependent on a complex interaction between the rhizobia and the host. It is initiated by flavonoid production by the host which induces nodulation factor production in the rhizobium. The bacteria then invade the host through a series of distinct processes resulting in the formation of an infection thread. Infection thread elongation continues until the bacteria reach the desired root tissue. The bacteria are then encapsulated by a host cell and differentiation into a nitrogen fixing bacteroid is completed.

1.5.1 Sinorhizobium meliloti symbiosis

S. meliloti 1021 forms a nitrogen fixing symbiosis with the leguminous plants *Medicago sativa* and *Medicago truncatula*. The bacterium is stimulated into forming this relationship through the uptake and subsequent transcriptional response caused by flavonoids produced by the plants. The flavonoid luteolin produced by *M. sativa* has been identified as one of the effectors to which *S. meliloti* responds. The luteolin molecule interacts with the transcriptional regulator NodD1 which in turn activates transcription of the *nod* genes. The two other Nod proteins encoded in *S. meliloti*, NodD2 which responds to an unknown effector molecule and NodD3 which is under the control of a complex regulatory cascade, can also initiate symbiosis. Each of the NodD proteins recognises the nod box promoter that allows them to activate the *nod* genes (Barnett and Fisher 2006; Perret, Staehelin and Broughton 2000).

The *nod* genes encode for proteins responsible for the assembly of the Nod factors. The Nod factors comprise a backbone of β-1,4-linked N-acetyl-D-glucosamine polymer that can vary in their length inter and intra species. This N-acetyl-glucosamine backbone is in turn acetylated with fatty acid moieties of various chain lengths. The NodABC proteins act to construct the core of the factor with numerous other Nod related proteins acting to tailor the backbone structure to impart host strain characteristics (Geurts, Fedorova and Bisseling 2005; Oldroyd and Downie 2004).

The Nod Factors are released by the bacteria and induce a number of physiological changes in the root. These responses include a spike in Ca^{2+} levels in the root hairs which is followed by root curling which traps the bacteria or single bacterium in a tight

colonised curled root hair (CCRH). In addition to these events root cortex cells are stimulated to initiate mitosis which give rise to the nodule primordium that receives the invading bacteria (Murray 2011).

The next stage of the infection process is the internalisation of the bacteria and the formation of the infection thread. Bacteria in the CCRH continue to produce Nod Factors and start to produce exopolysaccharide which induces an ingrowth into the root hair which results in bacteria gaining access to the interior plant tissue. The bacteria form an infection thread that progresses through the plant. Only bacteria at the tip of the infection thread are actively dividing with the remainder of the bacteria in the thread senescing (Gage 2004; Gage 2002).

The infection thread undergoes successive rounds of formation to break through each layer of tissue until it reaches the inner plant cortex. At this point bacteria escape the infection thread into a plant cell and in the process becomes encapsulated by an unwalled membrane. This entire intracellular structure is referred to as the symbiosome which is thought to be a derivative of the lytic vacuole. Prior to forming nitrogen fixing bacteroids the entire symbiosome divides as if it was a single organism (Brewin 2004; Robertson and Lyttleton 1984).

Symbiosome survival at this point is essential for the formation of a fully functional mature nitrogen fixing nodule. *S. meliloti* protects itself from the plant defences through the production of lipopolysaccharide with mutants defective in its production succumbing to cell lysis (Glazebrook, Ichige and Walker 1993). The plant also supports the intracellular symbiosome by providing essential nutrients and key conditions required for differentiation into bacteroids and nitrogen fixing activities. The conditions inside the symbiosome are microoxic and bacteria that complete differentiation into bacteroids are now in a position to begin nitrogen fixation. A low oxygen environment is essential for nitrogen fixing and the genes that encode the nitrogen fixing proteins are under the control of an oxygen sensing regulatory cascade which also activates the enzymes for microoxic respiration (Fischer 1994). The regulators involved in this cascade are the FixL, FixJ, FixK, NifA and a σ^{54} that control the *fix* and *nif* genes that act together to fix nitrogen. The expression of these genes is inversely linked to general metabolic processes in the newly formed bacteroid. The fixation process results in 16 molecules of ATP and 8 electrons providing the approximate energy required to reduce

1 molecule of N_2 to two molecules of NH_4 (Barnett and Fisher 2006; Oldroyd and Downie 2004).

The action of reducing N_2 to NH_4 is a highly energetic process and is supported by the plant. The plant provides a carbon source in the form of malate and also through the activity of sucrose synthase that catabolises sucrose into UDP-glucose and fructose which is metabolised by the plant into malate prior to transport to the bacteroid (Poole and Allaway 2000). Also essential to the continued fixation activity is the presence of a microoxic environment which is maintained by the plant through the production of legheamoglobin which sequester free oxygen and tightly control the oxygen status of the bacteroid (Ott et al. 2005).

1.6 Iron acquisition in free living Rhizobia

The α-proteobacteria is a taxonomic group of bacteria which includes a number of important genera, including some which form a relationship with higher eukaryotes either symbiotically or through pathogenesis. The *Rhizobiales* are a member of this phylum and exist as free living organisms in the soil or in a symbiotic nitrogen-fixing relationship with a leguminous host. Other notable members of the α-proteobacteria include the plant pathogen *Agrobacterium* and mammalian pathogens *Brucella* and *Bartonella*. Iron acquisition has been studied in detail in three of the rhizobial species namely, *S. meliloti*, *R. leguminosarum* and *B. japonicum.*

1.6.1 *Sinorhizobium meliloti*

S. meliloti is a Gram-negative bacterium that can be found either free living in the soil or in symbiosis with leguminous plants especially members of *Megicago spp*. The *S. meliloti* 1021 genome consists of three distinct replicons, a 3.7 Mb chromosome, the 1.4 Mb pSymA megaplasmid and the 1.7 Mb pSymB megaplasmid. The *S. meliloti* genome has been sequenced providing vital insights into the molecular processes relied on by the organism (Barnett et al. 2001; Capela et al. 2001; Finan et al. 2001; Galibert et al. 2001).

The primary method by which *S. meliloti* 1021 acquires iron is through the production of the citrate-dihydroxamate siderophore rhizobactin 1021. The structure of this siderophore is given in figure 1.6 and is similarly to other citrate-dihydroxamate siderophores such as aerobactin and schizokinen. Rhizobactin 1021 was first described by Persmark et al. (1993) and was the first known example of a terrestrial siderophore that possessed a lipid moiety, a trait associated with marine bacteria at that time. The genes that encode for the biosynthesis of rhizobactin 1021 are present on the pSymA megaplasmid and consist of a six gene operon transcribed from the positive strand named *rhbABCDEF*. Downstream from this operon is a transcriptional regulator *rhrA* transcribed from the negative strand and the gene encoding the cognate TBDT protein, *rhtA* for rhizobactin 1021 transcribed from the positive strand. Deactivation of this gene results in the abolition of rhizobactin 1021 acquisition. Analysis of the presence of transcripts for the *rhbA* and *rhbF* genes in mature alfalfa nodules showed that they are not expressed indicating that rhizobactin 1021 is not synthesised in the nodule (Lynch et

al. 2001). Rhizobactin 1021 is transported across the inner membrane by the single unite transporter, RhtX. This protein is transcribed upstream of the rhizobactin 1021 biosynthesis operon and is transcribed on the positive strand. It is also transcriptionally coupled to the *rhbABCDEF* operon. It can transport both rhizobactin 1021 and the structurally similar siderophore schizokinen but is unable to transport aerobactin which is of a slightly different core structure (O Cuiv et al. 2004). Figure 1.14 is a diagrammatic representation of the rhizobactin regulon as described above.

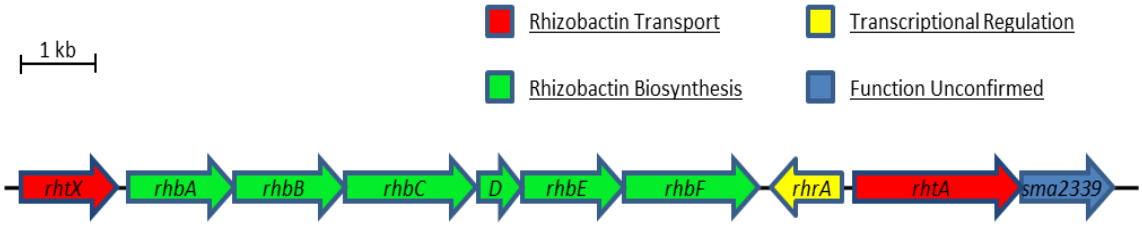


Figure 1.14: Representation of the rhizobactin 1021 regulon

The activity of the AraC-like regulator RhrA has been assessed by mutagenesis and RNase protection assays which demonstrated that it functions to recognise regions upstream of rhizobactin 1021 transport and biosynthesis (Lynch et al. 2001). There was no expression of the *rhbABCDEF* and *rhtA* genes in an *rhrA* mutant which confirmed that RhrA is the transcriptional activator of the rhizobactin 1021 associated genes. The general regulator of iron homeostasis in *S. meliloti* is RirA which was discovered by mutagenesis by Chao et al. (2005) and Viguier et al. (2005). RirA was shown by qRT-PCR to regulate the expression of *rhtA* and *rhbABCDEF* in an iron responsive manner and RirA was also shown to repress *rhrA* in iron replete conditions. The repression of *rhrA* by RirA demonstrates that the rhizobactin regulon is under dual control involving a transcriptional repressor and a transcriptional activator. The haem receptor *shmR* was also shown to be controlled be RirA by qRT-PCR analysis which showed that RirA acted in a global manner. Indeed whole genome microarray experiments confirmed that RirA is active in a global manner Chao et al. (2005). Importantly the *rirA* mutant showed no defect in symbiotic nitrogen fixation (Viguier et al. 2005).

In addition to rhizobactin 1021 production *S. meliloti* engages in xenosiderophore utilisation to satisfy its iron requirement. Characterisation of transport systems involved in siderophore utilisation identified the TBDT proteins FhuA, FoxA and ShmR responsible for the outer membrane recognition of ferrichrome, ferrioxamine B and haem/haemoglobin respectively (O Cuiv et al. 2008; Battistoni et al. 2002). The inner

membrane transport system of these siderophores display an interesting shared mechanism. The hydroxamate siderophores ferrichrome and ferrioxamine B are recognised by the periplasmic binding protein FhuP with haem being recognised by the periplasmic binding protein HmuT. These two PBPs interact with the same inner membrane complex, HmuUV the permease and ATPase respectively. The shared transport system for uptake of haem, ferrichrome and ferrioxamine B is shown in figure 1.15.

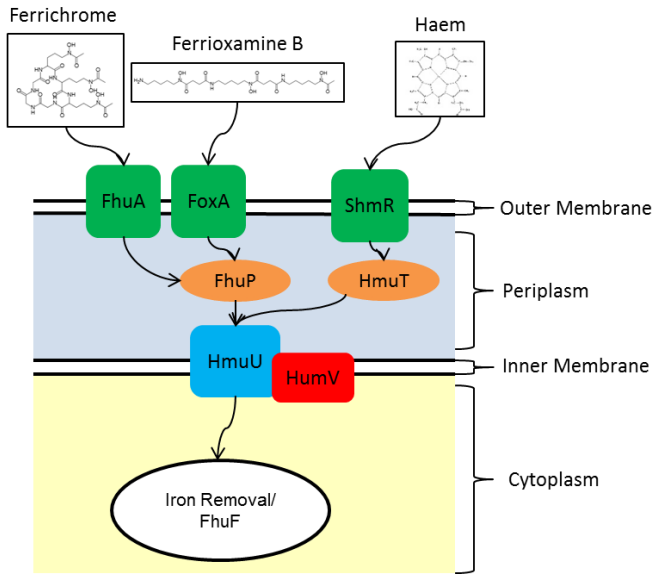


Figure 1.15: Representation of haem and xenosiderophore uptake by *S. meliloti*

1.6.2 *Rhizobium leguminosarum*

R. leguminosarum is similar to *S. meliloti* as it is a Gram-negative bacteria that is found in either a free living state in the soil or in symbiosis with specific leguminous plants. The host legume for the symbiosis varies depending on the particular biovar of *R. leguminosarum* strain. The genome of *R. leguminosarum* consists of seven replicons formed by a 5 Mb chromosome and six independent megaplasmids ranging in size from 800 kb to 150 kb. In response to iron stress encountered in the free living state *R. leguminosarum* produces the siderophore vicibactin which has biosynthesis genes located on the pRL12 plasmid. This siderophore has a cyclic trihydroxamate structure that is constructed from D-3-hydroxybutanoic acid and N_2-acetyl-N_5-hydroxy-D-ornithine assembled through a series of ester and peptide bonds (Dilworth et al. 1998). The biosynthesis proteins for vicibactin are transcribed from four proximally located

operons that contain a total of eight genes, *vbsGSO, vbsADL*, *vbsC* and *vpsP*. Based on mutagenesis and *in silico* analysis a biosynthesis pathway has been proposed which is based around the NRPS protein VbsS (Carter et al. 2002).

An outer membrane receptor for vicibactin, FhuA1 has been identified and is genetically located proximal to the siderophore biosynthesis genes and shows strong homology to the ferrichrome receptor, FhuA in *E. coli*. There is also a homolog of a siderophore reductase located downstream of the receptor FhuA1 possibly involved in iron release off the siderophore. In addition to this FhuA homolog there is a second FhuA homolog that is encoded proximal and divergently to genes homologous to *fhuBCD* for hydroxamate siderophore transport in *E. coli*. However this second homolog has been shown to be a pseudogene with no activity with the *fhuBCD* genes encoding for vicibactin inner membrane transport (Yeoman et al. 2000; Stevens et al. 1999). *R. leguminosarum* can also utilise haem as an iron source. The functions for haem acquisition are provided by the proteins ShmR and HmuTUV in a similar arrangement to that found in *S. meliloti* but lacking the interplay with hydroxamate siderophores. The *shmR* gene encodes for the outer membrane receptor and is located as a single transcript while the *hmuTUV* genes are located distally and encode for inner membrane transport. As with other organisms the uptake of haem and siderophores is dependent on the presence of a functional TonB protein (Wexler et al. 2001).

Iron homeostasis in *R. leguminosarum* is under the dual control of the two transcriptional regulators RirA and Irr. These two proteins act in unison with one another with RirA repressing its target genes expression under iron replete conditions and Irr repressing gene expression under iron deplete conditions. RirA functions to sense iron through the presence of Fe-S clusters in the cell while Irr functions by sensing the oxidation status to cellular haem. However Irr in *R. leguminosarum* has a less complex relationship with haem than Irr from *B. japonicum*. Haem acts to block Irr from binding DNA in *R. leguminosarum* where in *B. japonicum* haem results in the degradation of the Irr protein in a negative feedback loop associated with the cellular iron status (White et al. 2011; Todd et al. 2006; Todd et al. 2002).

1.6.3 *Bradyrhizobium japonicum*

B. japonicum occupies a similar environmental niche as *S. meliloti* and *R. leguminosarum* in that it resides as both a free living and symbiotic organism. Its genome is comprised of a single 9 Mb chromosome and unusually amongst the rhizobia contains no megaplasmids (O'Brian and Fabiano 2010). There is no siderophore reported to be produced by *B. japonicum* but it does encode for a number of TBDT proteins, eleven in total, five of which have being shown to be regulated by the Irr protein. Of the remaining six receptors two are proposed to be involved in cobalamin and nickel uptake with no functions being assigned to the remaining four. Perhaps these genes require their cognate substrate for them to be transcriptionally actively such as the previously described examples of *fptA* and *fecA*. However this is only speculative observation. The five receptors that are strongly activated in iron deplete conditions were identified by analysis carried out by (Small et al. 2009; Yang et al. 2006). To date three of the TBDT proteins have been functionally characterised, FhuE, PyoR and FegA which are responsible for rhodotorulic acid, pyoverdine PL-8 and ferrichrome uptake respectively (Small 2011; Benson, Boncompagni and Guerinot 2005). The FegA protein is homologous to the ferrichrome receptor FhuA in *E. coli* and was shown to be essential for ferrichrome acquisition in *B. japonicum*. Interestingly a *fegA* mutant showed a highly defective phenotype when assessed for nitrogen fixation which demonstrates that this receptor has activity beyond that of ferrichrome transport as ferrichrome would not be present in a nodule (Benson, Boncompagni and Guerinot 2005). Further studies carried out in a *B. japonicum* LO derived strain into the effect of the *fegA* mutation on nitrogen fixation showed that the deficient phenotype was specific to the strain *B. japonicum* 61A152 (Small et al. 2009). Analysis of *B. japonicum* for the use of other xenosiderophores demonstrated that it can utilise ferrioxamine B in addition to ferrichrome, rhodotorulic acid and pyoverdine PL-8 but the receptor involved in its transport has not been characterised (Benson, Boncompagni and Guerinot 2005; Plessner, Klapatch and Guerinot 1993). Indeed the transport of rhodotorulic acid may suggest that *B. japonicum* can utilise coprogen also as the transport of these two siderophores has been shown to be overlapping (Hantke 1983). Interestingly there are no ABC transporter systems transcribed adjacent to any of the TBDT leading to an incomplete understanding of xenosiderophore utilisation *B. japonicum* (O'Brian and Fabiano 2010). The only inner membrane protein shown to be necessary for ferrichrome utilisation is FegB, which is transcribed downstream of *fegA.*

This is an unusual protein that bears no resemblance to known siderophore transporters but is conserved across many species as reported by Benson, Boncompagni and Guerinot (2005) however as observed by Small et al. (2009) when assessing ferrichrome uptake by *B. japonicum* LO *fegB* was found to be absent but ferrichrome was still utilised. The authors propose that FegB could assemble into a transporter and function similarly to the rhizobactin 1021 inner membrane MFS, RhtX but also state that this is unlikely due to the low similarity between the two proteins. Unpublished observations by Cooke and O'Connell, where BLASTp analysis of FegB was carried out showed similarity to putative peptidases in many strains. This coupled to the knowledge that FegB is known to associate with the inner membrane opens up the possibility that it could act to degrade the amine bonds in ferrichrome therefore releasing the iron into the periplasm. If this was the case FegB may be a new class of protein involved in siderophore uptake.

In addition to xenosiderophore uptake, *B. japonicum* also acquires exogenous haem from the environment via the HmuR TBDT with the assistance of HmuPTUV comprising a transcriptional modulator of *hmuR* designated HmuP and an ABC transporter system HmuTUV (Escamilla-Hernandez and O'Brian 2012; Nienaber, Hennecke and Fischer 2001). Interestingly the genes for haem uptake are co-located in *B. japonicum* which is not the case in *S. meliloti* and *R. leguminosarum*. There are similarities in the haem uptake in all three strains however, in that each relies on a TBDT protein but has a dispensable associated ABC transporter system, *hmuTUV*. Mutagenesis of any of the genes *hmuTUV* results in a reduction of haem utilisation, not abolition, indicating a second transporter system that is yet to be discovered in each strain (Battistoni et al. 2002; Nienaber, Hennecke and Fischer 2001; Wexler et al. 2001). Perhaps a derivative of the secondary haem transporter system found in *E. coli* could be fulfilling this role as the DppA dipeptide transport system was found to recognise haem in the absence of the Hmu inner membrane transport system (Letoffe, Delepelaire and Wandersman 2006). There are multiple dipeptide transport systems present in these three rhizobial strains which leaves many possibilities for the second transport system if the phenomenon of dipeptide systems compensating haem uptake is widespread.

1.7 Mass-Spectrometry analysis of siderophores

An essential aspect of the experimental work described as part of this document is the electrospray ionisation mass spectrometer analysis of the siderophore rhizobactin 1021. This technique has proven to be essential to the analysis of many biologically relevant substrates from small metabolites to full proteins, providing accurate molecular weights with the inference of additional structural information from the production of collision induced fragmentation of the parent ion.

The process of producing electrospray ions can be split into two distinct events; first the sample is dispersed as highly charged droplets at near atmospheric pressure followed by evaporation of these droplets allowing the contained ions to be analysed. Sample dispersal and charging is often achieved via a nebulizer linked to a capillary tube. The nebulizer forms the aerosol of droplets that are passed through the capillary which maintains a high electric field resulting in a charge being imparted onto each droplet. This droplet is maintained by a high pressure gas flow through the capillary. Evaporation of the charged droplet is essential for the detection of the desired molecular ions and is accomplished by the addition of heat to the capillary which progressively reduces the droplet until the electrical Coulomb forces reach a level equivocal to the surface tension of the droplet. When this point is reached droplet fission occurs resulting in the dispersal of the contained ions onto the detector. As this occurs in an electrical field the ions must be charged in order for them to interact with the electric field. Depending on the electron preference of the molecule the instrument can detect in the negative or positive modes. The presence of an electron acceptor such as oxygen results in best detection in the negative mode and the presence of positively charged groups results in best detection in the positive mode. A charge is imparted onto molecules through the addition of a protonation source such as acetic acid (Smith et al. 1990).

The analysis of siderophores via ESI-MS has allowed the accurate measurement of molecular weights and the production of signature fragmentation patterns. There are common themes to the spectra observed while analysing substrates by ESI-MS and the analysis of siderophores is no different. The most commonly encountered features are:

- Contaminants can be present in the sample. This is a common occurrence especially when working with samples extracted from complex mixes such as microbial supernatants and media constituents
- Degradation can occur of the target molecule due to storage conditions, sample preparation or inherent instability of the compound. It has been noted that ligand free rhizobactin 1021 can slowly dehydrate in a time dependent manner which is represented by a loss of 18 Da by mass spectra analysis (Persmark et al. 1993).
- Due to the relatively benign conditions under which the sample is ionised molecules that have a propensity to form oligomers can be detected by the presence of spectral peaks at positions M, M+M, M+M(n), where M is the mass of the molecule and n is any number. This was observed for rhodotorulic acid where both dimers and iron complexes were observed (Gledhill 2001).
- The formation of metal adducts with the target molecule is also a common occurrence. Sodium, potassium and other metals can associate with the target molecule resulting in peak shifts to larger sizes on the mass spectrum. This has been observed in an analysis of schizokinen where the iron bound siderophore was found to have a mass 23 Da heavier than expected representing a sodium adduct (Storey et al. 2006).
- Small variants can occur caused by the presence of one extra H^+ ion due to the protonation of the sample. Also hydrogens can be lost and replaced by metals either in the form of bound iron for siderophores and/or including metal adducts. These forms were also observed by (Storey et al. 2006) where the ferrated siderophore was represented by [M]-[3H]+[Fe]+[H]. The first three hydrogens are lost to accommodate the iron molecule with the last hydrogen representing the protonating hydrogen.

With awareness of the inherent variability in the ESI-MS method this method can be used to extract exact masses for analytes present in a sample. This along with accurate fragmentation patterns allows for any changes in siderophore structure to be identified which is of paramount importance in the proceeding experimental section.

1.8 Summary

Bacteria have evolved complex and elegant strategies to address the physiological stress related to iron insolubility with these methods also addressing stress related to toxicity by strictly controlling intracellular iron levels as discussed previously. Many of these mechanisms involve the synthesis and secretion of powerful iron chelators but many other weaker iron chelators such as citrate have been exploited for their chelated iron. Perhaps unsurprisingly microbes have also developed means to extract iron from storage molecules and proteins such as haem, haemoglobin and transferrin to name a few. Taken together all these systems act to ensure that bacteria can acquire sufficient iron to grow and thrive in its environmental niche which can range from highly nutritional locations such as the gut to more competitive areas such as soil and minimal nutritional locations such as marine environments. To ensure that the array of iron acquisition methods function efficiently a complex network of regulatory elements control iron acquisition, all of which are reliant on the iron status of the cell either directly or indirectly.

The experimental work described herein centres on the siderophore related iron acquisition methods of the rhizobial species *S. meliloti* 2011. The primary focus of this work was to assess the unresolved questions in relation to the biosynthesis of the endogenous siderophore rhizobactin 1021. The Sma2339 protein, named RhbG by many publications has been speculated to perform the role of lipid addition to rhizobactin 1021 based on genome location and homology to other siderophore biosynthesis proteins (Miethke et al. 2011; Challis 2005; Viguier et al. 2005). However no experimental evidence for this role is presented in the literature which leaves its role open to question. The experimental analysis presented here addresses the uncertainty surrounding the role of *sma2339* in rhizobactin 1021 biosynthesis. It was found that *sma2339* is not responsible for the addition of the lipid moiety to rhizobactin 1021. Characterisation of a *S. meliloti* 2011*sma2339* mutant confirms that it is vital to the efficient production of rhizobactin 1021 as a drastic reduction in siderophore production occurs in its absence as quantified by the liquid Chrome Azurol Sulphonate assay.

In addition to analysis of rhizobactin 1021 biosynthesis, xenosiderophore acquisition in *S. meliloti* 2011 was investigated. Previous studies have shown that *S. meliloti* 2011 can utilise the siderophores ferrichrome, ferrioxamine B along with other siderophores

structurally related to rhizobactin 1021. It was found that the Sma1747 protein, now suggested to be renamed FhuE is the cognate TBDT protein for the fungal siderophore coprogen. This was confirmed through complementation and mutagenesis studies in *S. meliloti* 2011 derived strains and in the *S. meliloit* 2011 wild type strain. The exact mechanism for inner membrane transport of coprogen remains elusive as uptake still occurs in the absence of a putative inner membrane transport system located proximal to *fhuE*. The possibility of a similar situation to that in *E. coli*, where a number of hydroxamate siderophores are transported through the PBP FhuD, was also assessed in *S. meliloti*. FhuP of *S. meliloti* which is involved in the transport of the hydroxamate siderophores ferrichrome and ferrioxamine B was the foremost candidate for this role as FhuD from *E. coli* also interacts with these siderophores along with coprogen (Braun, Hantke and Koster 1998). However, a *S. meliloti* Rm818*fhuP* mutant strain complemented with *fhuE in trans* showed no disruption of coprogen utilisation indicating that *fhuP* is not a member of the redundant system.

A prominent feature of siderophore uptake in *S. meliloti* 2011 is the presence of AraC-like genes proximal to genes encoding TBDT proteins. Here we investigate the possible regulation of the newly identified *fhuE* receptor gene by the presence of its cognate substrate coprogen, as an AraC-like gene *sma1749* is transcribed proximal to the receptor. This analysis demonstrated a clear response due to the presence of coprogen under iron deplete conditions. In addition, the role of ferrichrome in the control of FhuA expression was also investigated. The *fhuA* gene is present in an unusual arrangement with a proximal AraC-like gene, *smc01610* as they appear to be co-transcribed. This analysis also demonstrated induction of the receptor under iron limited conditions in the presence of ferrichrome in an apparent RirA independent fashion.

The study of *S. meliloti* and indeed other rhizobial strains as model organisms for iron acquisition has allowed for the identification of a number of key variations on the main themes of iron acquisition in Gram-negative bacteria. Perhaps most notable is the discovery of the RirA and Irr regulatory mechanisms that were first described in the rhizobia and a homolog to Irr was subsequently found to be regulating iron responsiveness in the bovine and human pathogen *Brucella abortus* (Anderson et al. 2011). Also, both RirA and Irr were identified as the antiparallel regulators of iron homeostasis in the widespread plant pathogen *A. tumefaciens* (Hibbing and Fuqua 2011). As Fur is the regulator of iron homeostasis in other model organisms such as *E.*

coli and *P. aeruginosa* this demonstrates the importance of having a number of model organisms as RirA and Irr are not of significance to iron acquisition in these organisms (McHugh et al. 2003, Vasil and Ochsner 1999).

The identification of FhuE provides further insight into the iron acquisition strategies employed by *S. meliloti* in a free living state. Further to this, expression analysis showing the siderophores coprogen and ferrichrome acting as effectors for the expression of *fhuE* and *fhuA* respectively represents a finely balanced technique employed to conserve energy. Also this type of regulation allows *S. meliloti* to evade detrimental compounds present in the extracellular milieu such as phage and toxins released from neighbouring organisms.

The work presented herein also constitutes a significant clarification on the biosynthesis pathway for rhizobactin 1021. As the role of Sma2339 in rhizobactin 1021 biosynthesis is not lipid addition this allows for the consideration of other mechanisms by which this is achieved. Indeed this may contribute to more widespread cellular functions being considered for the role of lipid addition such as a periphery protein to the fatty acid biosynthesis pathway as the phenomenon of siderophore acylation is vaster than the distribution of Sma2339 homologs. Characterisation of other mechanisms was beyond the scope of the analysis described in this thesis but the possible alternatives will be discussed in summation.

The variety of siderophores displaying lipid structures continues to increase in number. However, the mechanism behind how these structures are attached to the siderophore has yet to be elucidated. The study of rhizobactin 1021 biosynthesis in *S. meliloti* presents an opportunity to investigate siderophore acylation in a directed manner for a number of reasons; a detailed genome sequence in available allowing for genome manipulation to be carried out relatively simply, *S. meliloti* has been well characterised with regard to iron acquisition and also numerous techniques for mutagenesis and complementation have been identified that work efficiently. Identification of a *S. meliloti* mutant strain producing non-acylated rhizobactin 1021 would provide an excellent reference point from which the mechanism of acylation of many other siderophores could be elucidated.

Chapter Two

Methods and Materials

2.1 Bacterial strains, Primer sequences and Plasmids

Table 2.1: Bacterial Strains

Strain	Genotype	Source
***E. coli* JM109**	*e14*$^-$(*mcr*A$^-$), *rec*A1, *end*A1, *gyr*A96, *thi*-1,*hsd*R17($r_K^-m_K^-$), *sup*E44, *rel*A1, λ^-,Δ(lac-proAB), [F'*tra*D36, *pro*AB, *lac*I^{q}ZΔ M15]	Promega
***S. meliloti* 2011**	Wild type, Nod$^+$, Fix$^+$, Sm^R	(Meade et al. 1982)
S. meliloti* 2011*rhtX-3	Cm^R in *rhtX*	(O Cuiv et al. 2007)
S. meliloti* 2011*rhbA62	Km^R in *rhbA*	(Lynch et al. 2001)
S. meliloti* 2011*sma2339	Km^R replacement of *sma2339*	This Study
S. meliloti* 2011*rhbD	Tc^R in *rhbD*	This Study
S. meliloti* 2011*sma2339rhbD	Km^R replacement of *sma2339,* Tc^R in *rhbD*	This Study
S. meliloti* 2011 *rhtX-3sma1747	Cm^R in *rhtX*, Km^R in *sma1747*	This Study
***S. meliloti* Rm818**	pSymA cured strain	Dr. Michael F. Hynes, Calgary
S. meliloti* Rm818*fhuP	pSymA cured strain, Km^R in *fhuP*	This Study

Table 2.2: Plasmids

Plasmid	Description	Source
pJQ200sk	Suicide vector, Gm^R, *mob*, *sacB* from *Bacillus subtilis*, p15a origin	(Quandt and Hynes, 1993)
pRK600	Provides transfer functions, Cm^R	(Finan *et al*.,1986)
pBBR1MCS-5	Broad host range cloning vector, Gm^R, *mob*	(Kovach *et al*., 1995)
pRK415	Broad host range cloning vector, Tc^R, RK2 derivative	(Keen et al. 1988)
pUC4K	Source of Km^R cassette, Amp^R	Amersham Pharmacia
pHP45-ΩTc	Source of Tc^R cassette, Amp^R	(Kovach *et al*., 1995)

Table 2.3: pJQ200sk Derived Plasmids

Plasmid	Description	Source
pJQ2339F12Km	Contains region for mutagenesis of *sma2339*	This Study
pJQrhbDF12Tc	Contains region for mutagenesis of *rhbD*	This Study
pJQ1747F12Km	Contains region for mutagenesis of *sma1747*	This Study
pDK2.0K A/S	Contains region for mutagenesis of *fhuP*	(O Cuiv et al. 2008)

Table 2.4:pBBR1MCS-5 and pRK415 Derived Plasmids

Plasmid	Description	Source
pBBR1MCS-5 Derived Plasmids		
pCC101	Contains *rhtAsma2339* for expression	This Study
pCC102	Contains *rhbDEF* for expression	This Study
pCC103	Contains *rhbEF* for expression	This Study
pKC102	Contains *sma1746-41* for expression	This Study
pDK104	Contains *fhuP* for expression	(O Cuiv et al. 2008)
pRK415 Derived Plasmids		
pKC101	Contains *sma1747* for expression	This Study

Table 2.5: Primer sequences for cassette mutagenesis

Name	Sequence 5` → 3`
Sma2339 Del1F	CACCACGGGCCCCCGGTCTTTATCCATCCATCG
Sma2339 Del1R	CACCACGGATCCTTAAAAAACCTTTGTCAGCGAGAC
Sma2339 Del2F	CACCACGGATCCAGCGGCCGACCGCACCAGTT
Sma2339 Del2R	CACCACGAGCTCCCCGCTCGGCTGTCAGCCAA
Sma2339/Km For	GGCGGCCTAGCCGACGAGTCG
Sma2339/Km Rev	ACTGAGAAGCTGGGCGGCAGG
rhbDmut1F	CACCACGGGCCCGGCGAGGCCGGCAAGCCATGG
rhbDmut1R	GAGGAGGTCGACCTCAAGCATGTGCCAGAACCTTTTCCG
rhbDmut2F	CACCACGTCGACGAGGAGCGAGCCGGACACTGCC
rhbDmut2R	CACCACTCTAGAACCACTTCCTCGCCGAAGCGGCACG
rhbDTc For	ATGCTGGCGCCGACGCTGC
rhbDTc Rev	CATCTTCCATTGCGGCACG
TetCas For	CTAGTAACGGCCTCGAGTGTGCTGG
TetCas Rev	CCGCCAGTGTGCTCGAGATCTGCAGAA
1747mut1F	ATCTCTAGGGCCCTTGGGCTGCACAGCATTT
1747mut1R	ATCTCTAGTCGACGCCACGGTCGAGCCAT
1747mut2F	ATCTCTAGTCGACGGGTCGCTATTTCTTCGG
1747mut2R	ATCTCTAGGATCCCGCGACGTGTTTCTCGTC
47kmscreenF	CGGCTTGAGTTTCAGTTCCAACG
47kmscreenR	CTTTCCACATTGTCGAGACTTGC
FhuPKmF	TTGTCTCGTCTGCTGCTCTGCCGCC
FhuPKmR	TCATTTTCTTCAAGCTTATCCAATAGG

Table 2.6: Primer sequences for cloning

Name	Sequence
rhtA/sma2339 For	CACCACGGTACCTCTGGTGCGCAGGGGGCG
rhtA/sma2339 Rev	CACCACGAGCTCTCATGAGTTGCTCCTGTCGGGATCC
rhbDEF For	CACCACGGGCCCGGAAAAGGTTCTGGCACATGCTTGAG
rhbEF For	CACCACGGGCCCTTCATTCAACAATTCGGATTATGATCATGA
rhbF Rev	CACCACTCTAGATTACGCGGCGTTGCTGCGCTTGTCG
Sma1747F	ATGCCATTCGAATTCTACGTCAATTGGAACATGAAC
Sma1747R	ATGCCATTCGAATTCTTACCAACTCTGGCGCAGG
Sma1746-41F	ATGCCATTCGGGCCCAACGTTGAATGGAGTGTGGC
Sma1746-41R	ATGCCATTCAAGCTTCGCTCCTCTCCTTTTGTGG

Table 2.7: Primer sequences for qPCR

Name	Sequence	Target Gene
Smc03979 qFor2	GATCCACTCCTACACCAACG	*gapA*
Smc03979 qRev2	AGCTTGCCCTTGAGTTCC	
FhuA qFor1	TCCAGATAACCAAGCAGAACG	*fhuA*
FhuA qFor2	GGTATAGGAGAAGGCGCTTATG	
Smc01610 qFor1	CGCATGCAGAACAGGATCAG	*smc01610*
Smc01610 qRev1	AACCACACCATTCCACTCTC	
Sma1747 qFor1	GGGTAAGCAGTATGAAGTCGG	*sma1747*
Sma1747 qRev1	ATGTCTGACCTTTTCGACGG	
Sma1749 qFor1	TCATAAGGATCGGGTGTTGC	*sma1749*
Sma1749 qRev1	TGGTAGGGCTCTTTATTTCGTC	
FhuP qFor2	GGTCTTCGAGAGAGAAGAACAG	*fhuP*
FhuP qRev2	GCTTCCATGAGAACTTTCGG	

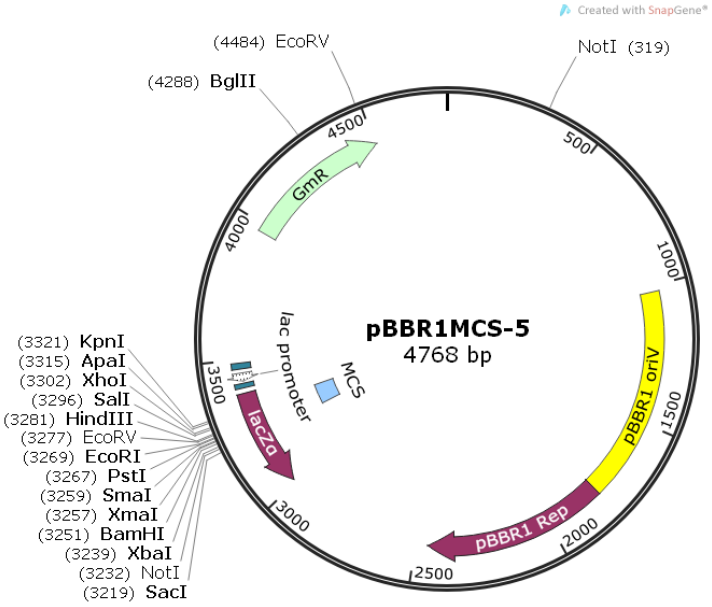


Figure 2.1: Vector map of pBBR1MCS-5

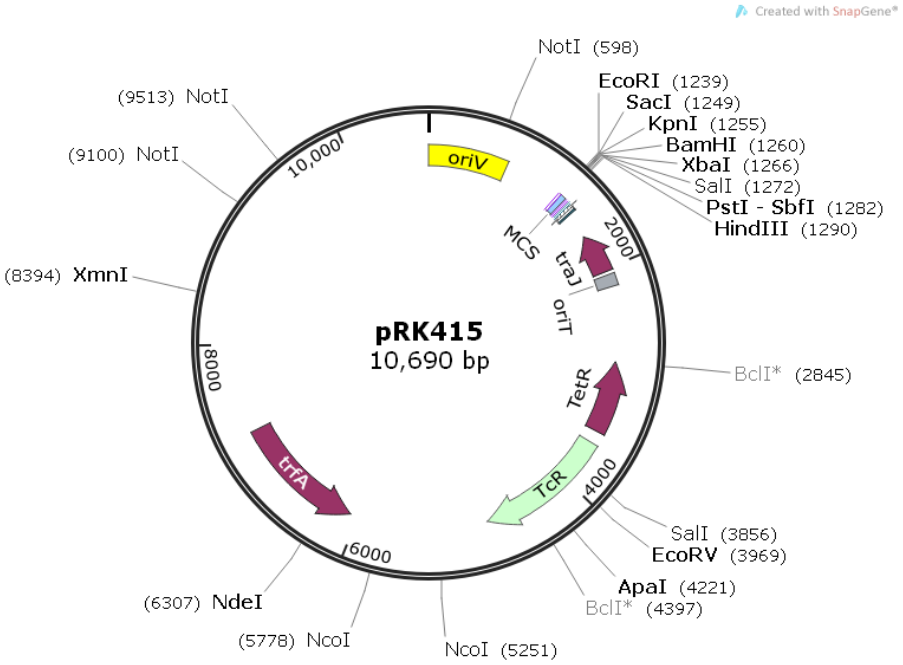


Figure 2.2: Vector map of pRK415

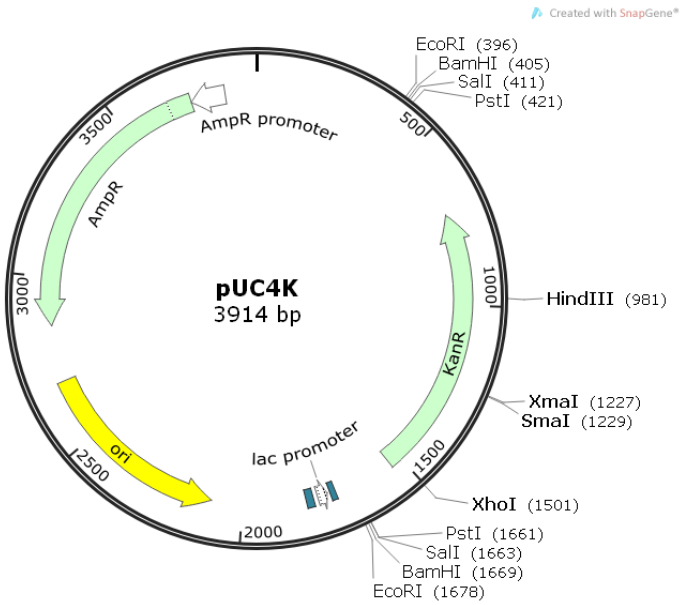


Figure 2.3: Vector map of pUC4K

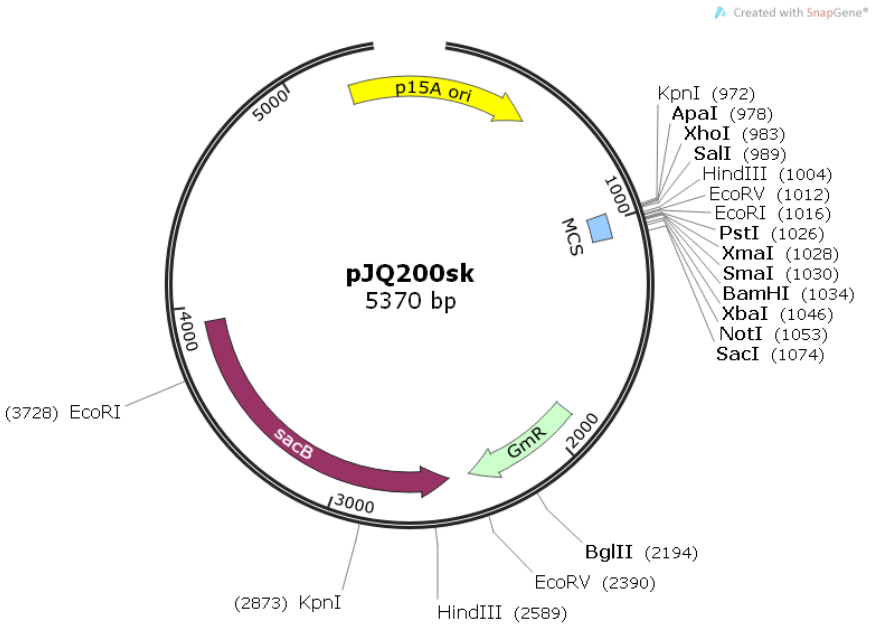


Figure 2.4: Vector map of pJQ200sk

2.2 Microbiological Media

Solid media contained 15 g/L Oxoid Bacteriological Agar. Other chemicals were from Sigma Aldrich and BDH Chemicals Ltd and of analar grade. Glassware used for low iron media was washed in 2 M HCl and rinsed in ultra-pure deionised water. Low iron medium was prepared in ultra-pure water and supplemented as desired. Distilled water was used to prepare complex media and sterilisation was achieved by autoclaving at 15 psi for 20 min.

- Preparation of H_2O for low iron solutions

A cation exchange column was prepared using Amberlite 200 resin and ultra-pure dH_2O. The Amberlite 200 resin is a strongly acidic cation exchange resin that efficiently chelates divalent metal ions. A reservoir of water was prepared and the inflow was mediated via a capillary tube. In order to prevent the column from running dry, the inflow and outflow rates were adjusted to equal rates. One column volume of water was allowed to flow through the column and was discarded. The column was then ready for the preparation of low iron water. Purified water was collected in a vessel that was washed thoroughly with 2 M HCl three times. Purified water was used to rinse residual 2 M HCl from the collection vessel.

- TY Broth

Used for routine culturing of rhizobia strains

Tryptone	5 g
Yeast Extract	3 g
$CaCl_2.2H_2O$	0.7 g

The pH was adjusted to pH 7.0 with NaOH and the volume was brought to 1 L with dH_2O. The solution was sterilised by autoclaving.

- Lysogeny Broth (LB)

Used for the culturing of *E. coli* and *P. aeruginosa* strains.

Tryptone	10 g
Yeast Extract	5 g
NaCl	10 g

The pH was adjusted to pH 7.0 with NaOH and the volume was brought to 1 L with dH_2O. The solution was sterilised by autoclaving.

- SOB Medium

Tryptone	20 g
Yeast Extract	5 g
NaCl	0.5 g
KCl	2.5 mM
dH_2O 1 L	
pH 7.0	

The solution was sterilised by autoclaving and allowed to cool to 55°C. Sterile solutions of $MgCl_2$ (1 M) and $MgSO_4$ (1 M) were added to final concentrations of 10 mM.

- SOC Medium

SOB broth was supplemented to a final concentration of 20 mM using a filter sterilised 50% glucose solution.

- SB Medium

Tryptone	30 g
Yeast Extract	20 g
MOPS	10 g

The solution was adjusted to pH 7.0 with NaOH brought to 1 L with dH_2O and sterilised by autoclaving. Once sterilised 20 ml of filter sterilised 20% (w/v) glucose and 10 ml of 1 M $MgCl_2$ was added.

- Tris Medium (Modified from Smith and Neilands, 1984)

Primarily used to illicit siderophore production.

Trizma Base	12.1 g
Glutamine	1 g
Succinic acid	1 g
KH_2PO_4	0.1 g
K_2HPO_4	0.1 g
$Na_2SO_4.10H_2O$	0.0317 g
Casamino Acids	3 g
H_2O	1 L
pH 6.8	

After autoclaving the medium was supplemented with 810 µl of 1 M $MgSO_4$ and 680 µl of 1 M $CaCl_2$. The following vitamins were added to a final concentration of 1 mg/L: Inositol, Thiamine, Pantothenate and Biotin. Pyridoxine-HCl was also added at a final concentration of 0.2 mg/L.

- Iron Deplete Conditions

To achieve an iron limited growth environment for *S. meliloti* 2011 2,2'dipyridyl was added to the medium at a final concentration of 300 µM after autoclaving. To achieve iron limitation in mutants lacking siderophore production/transport the 2,2'dipyridyl concentration was adjusted based on experimental conditions.

2.3 Solutions and Buffers for DNA analysis

- **General Buffers**

- TE Buffer

Tris-HCl	10 mM
Na_2-EDTA	1 mM
pH 8.0	

- PBS (10X)

Primarily used for suspension of siderophores and bacterial cells.

Na_2HPO_4	10.9 g
NaH_2PO_4	3.2 g
NaCl	90 g
Brought to 1 L with dH_2O	
pH is 7.2.	

- **Solutions for RbCl Competent Cells**

➢ TFB1 Buffer

RbCl	100 mM
$MnCl_2$	50 mM
Potassium Acetate	30 mM
$CaCl_2$	10 mM
Glycerol 15%	
pH 5.8	

The pH of the solution was adjusted with HCl before the $MnCl_2$ was added. The solution was then sterilised through a 0.22 μm sterile filter and stored at 4°C.

➢ TFB2 Buffer

MOPS	10 mM
RbCl	10 mM
$CaCl_2$	75 mM
Glycerol 15%	
pH 6.8	

The pH of the solution was adjusted with KOH. The solution was then sterilised through a 0.22 μm sterile filter and stored at 4°C.

- **Buffers for DNA analysis**

➢ 50X Tris Acetate EDTA (TAE) Buffer

EDTA	100 ml (0.5 M solution)
Glacial Acetic Acid	57.1 ml
Tris	242 g
dH_2O to 1 L	
pH 8.0	

The solution was diluted to 1X with dH_2O prior to use.

➢ 6X Gel loading Dye

Bromophenol Blue	0.25%
Xylene Cyanol	0.25%
Ficoll (Type 400)	15%

Made in dH_2O and sterilised by autoclaving.

Variations of the dyes were prepared as required. To allow for greater clarity of smaller DNA fragments bromophenol blue was not included as it can occlude faint bands. Similarly xylene cyanol was not included to increase the clarity of larger DNA fragments.

➢ DNA gel staining

DNA gel staining was achieved by steeping the gel in a 1 µg/ml ethidium bromide solution. Agarose gels for gel extraction were stained using 1X Gel Star gel stain incorporated into the gel. Gel Star was supplied as a 10,000X solution by Lonza and ethidium bromide was supplied at 10 mg/ml by Sigma Aldrich.

2.4 Antibiotics

Antibiotics were supplied by Sigma Aldrich and prepared as indicated by the supplier. Antibiotic stocks were prepared to suitable 1000X and 100X stocks and stored in the dark at -20 °C.

Ampicillin was prepared in dH_2O and used at a final concentration of 100 μg/ml in solid and liquid media for *E. coli*.

Chloramphenicol was prepared in 100% ethanol and used at a final concentration of 20 μg/ml in solid and liquid media for *E. coli*.

Gentamicin was prepared in dH_2O and was used at a final concentration of 20 μg/ml when culturing *E. coli* and *S. meliloti* strains.

Kanamycin was prepared in dH_2O and was used at a final concentration of 100 μg/ml in solid media and 50 μg/ml in liquid media for *S. meliloti*. For *E.coli*, kanamycin was used at a final concentration of 30 μg/ml in both solid and liquid media.

Streptomycin was prepared in dH_2O and used at a final concentration of 1 mg/ml for *S. meliloti*.

Tetracycline was prepared in 50% ethanol and was used at a final concentration of 10 μg/ml for *E. coli* and *S. meliloti* in solid and liquid media. Concentrated HCL (5-10 μl/ml) was added to increase solubility.

2.5 Isolation and Purification of Nucleic Acids

2.5.1 Plasmid DNA isolation using the Genelute Plasmid Miniprep Kit

The kit was used according to the manufacturer's instructions. A 1.5 ml aliquot of an overnight bacterial culture was pelleted by centrifugation at 13,000 rpm for 5 min. The supernatant was removed and the cell pellet was re-suspended in 200 µl of chilled re-suspension buffer. Cell lysis was achieved by the addition of 200 µl of lysis solution. The lysis was allowed to proceed for 5 min. After lysis, 350 µl of neutralisation buffer was added and mixed by inversion to precipitate the lysed cell debris. The precipitate was pelleted by centrifugation at 13,000 rpm for 10 min. While pelleting the precipitate a spin column was placed in a 2 ml centrifuge tube and 500 µl of column preparation solution passed through at 13,000 rpm for 1 min. The flow through was discarded. This step prepared the column for sample addition. Once the cell debris was pelleted the supernatant was transferred to the prepared spin column and centrifuged at 13,000 rpm for 30 sec. The flow through was discarded and 750 µl of wash solution was added and centrifuged at 13,000 rpm for 30 sec. The supernatant was discarded and the spin column dried by a further centrifugation at 13,000 rpm for 5 min. The spin column was placed in a fresh 1.5 ml tube and 100 µl TE buffer was added directly to the column membrane. This was allowed stand for 1 min prior to centrifugation at 13,000 rpm for 2 min. Plasmid DNA was stored at 4 °C or at -20 °C for long term storage.

2.5.2 Preparation of total genomic DNA using the Wizard Genomic DNA Kit

A 1.5 ml aliquot of an early stationary phase culture of *Sinorhizobium* spp. was pelleted in a microfuge tube at 13,000 rpm for 2 min. The supernatant was discarded and the cells re-suspended in 600 µl of the Nucleic Acid Lysis buffer. The cells were lysed by incubation at 80 °C for 5 min followed by cooling back to room temperature. To the lysate, 3 µl of RNase solution was added, inverted several times, and incubated at 37 °C for 60 min. To remove protein from the sample 200 µl of the Protein Precipitation solution was added and vortexed vigorously. This was incubated on ice for 5 min to aid precipitation and centrifuged at 13,000 rpm for 5 min. The supernatant was transferred to a fresh microfuge tube and 600 µl of phenol chloroform isoamyalcohol (25:24:1) was added and the suspension was mixed by gentle inversion. After centrifugation at 13,000 rpm for 5 min the aqueous phase was transferred to a fresh microfuge tube and

the phenol chloroform isoamylalcohol (25:24:1) step repeated. Phenol extraction was achieved through the addition of 700 µl of chloroform isoamylalcohol (24:1), mixed by inversion and centrifuged at 13,000 rpm for 5 min. The aqueous layer was transferred to a fresh microfuge tube and 600 µl of room temperature isopropanol was added. The sample was mixed gently until thread like strands of DNA became visible. The genomic DNA was collected at 13,000 rpm for 5 min. The pellet was washed twice with 500 µl of 75% ethanol, allowed to air dry and re-suspended in 100 µl of Rehydration Solution overnight at 4 °C. Genomic DNA was stored at 4 °C.

2.5.3 Purification of DNA using the illustra GFX PCR DNA and Gel Band Purification Kit

The kit was used in accordance with the manufacturer's specifications. Samples were prepared for binding to the spin column by the addition of 500 µl of Capture Buffer type 3 to 100 µl of sample and mixed by inversion. The Capture Buffer type 3 contains a pH indicator that remains a yellow/pale orange colour when the pH is optimal for binding. A dark pink/red colour is observed if the pH is too high and must be adjusted back to an optimal level using 3 M sodium acetate pH 5.0. Having ensured the pH is optimal for DNA binding to the column the sample was added to a spin column and placed in a 2 ml centrifuge tube. The tube was centrifuged at 13,000 rpm for 30 sec and the flow through discarded. Wash Buffer type 1 (500 µl) was added to the spin column and the microfuge tube was centrifuged at 13,000 rpm for 30 sec and the flow through discarded. The column was centrifuged for a further 5 min at 13,000 rpm to remove residual ethanol and then placed in a fresh 1.5 ml microfuge tube. Depending on the downstream applications of the purified DNA, the elution was achieved using either Elution buffer type 4 (TE buffer) or type 6 (DNase free H_2O). Elution buffer (30-50 µl) was added directly to the column capture membrane and allowed to absorb onto the membrane for 2 min at room temperature. Once absorbed the column was centrifuged at 13,000 rpm for 3 min. The eluted DNA was either used immediately or stored at -20 °C for later use.

DNA samples that required gel extraction were first subjected to agarose gel electrophoresis and stained using the Gel Star gel stain. The DNA fragment was visualised using a non-UV dark reader and the desired band excised using a scalpel. Up

to 300 mg of gel was placed in 500 µl of Capture Buffer type 3 and incubated at 55 °C for 15 min to melt the agarose gel. This sample was then added to a spin column in a 2 ml centrifuge tube and the procedure carried out as previously described.

2.5.4 Tri-reagent RNA extraction

A 15 ml aliquot was taken from *S. meliloti* cultures at OD_{600} 0.6 grown under conditions of interest and pelleted at 5,000 rpm for 5 min to harvest cells. The cells were suspended in 1 ml of PBS to wash away residual medium, transferred to a fresh 1.5 ml microfuge tube and pelleted at 6,000 rpm for 1 min. The PBS was removed and the cells were suspended in 1 ml of Tri-reagent and left stand at room temperature for 10 min to allow for complete cell lysis. (The procedure may be halted at this stage by storing the samples at -80 °C.) For every volume of Tri-reagent used, 0.2 volumes of chloroform was added. The samples were gently mixed by inversion until completely emulsified and allowed stand at room temperature for 10 min. The samples were spun at 13,000 rpm for 20 min at 4 °C to separate the aqueous phase (top phase) from the organic phase. The RNA is contained in the aqueous phase and this was removed to a fresh 1.5 ml microfuge tube. The RNA was precipitated through the addition of 500 µl of ice cold isopropanol per ml of Tri-reagent used and if desired, allowed stand for 15 min at -20 °C. The RNA was collected by centrifugation at 13,000 rpm for 15 min at 4 °C. The RNA pellet was washed 3 times with 1 ml of 75% v/v ethanol, mixed by inversion and spun at 13,000 rpm for 5 min at 4 °C after each wash. The ethanol was aspirated off and the pellet was left to air dry. Once dry the pellet was suspended in 20 µl RNase free water. Following suspension the $^{260}/_{230}$ and $^{260}/_{280}$ ratios were determined by nanodrop analysis to ensure purity. A DNase treatment was carried out for 45 min followed by a 10 min enzyme deactivation at 75°C. The RNA preparation was aliquoted and stored at -80°C.

2.5.5 cDNA Synthesis from extracted RNA

Before proceeding with cDNA synthesis it was ensured that all materials were RNase free and all H_2O was treated with DEPC. All cDNA synthesis was carried out using the Applied Biosystems High Capacity cDNA Reverse Transciption Kit and standard procedure.

Prior to cDNA synthesis the concentration of the test RNA was determined by nanodrop analysis. The RNA was brought to a concentration of 100 ng/µl in RNase free H_2O immediately prior to use. Each constituent of the reaction was thawed on ice and all, except for the Multiscribe™ Reverse Transcriptase were vortexed prior to use.

Table 2.8: Mastermix components for cDNA synthesis

Component	Volume (µl/Reaction)
10X RT Buffer	2.0
25X dNTP Mix (100mM)	0.8
10X RT Random Primers	2.0
Multiscribe™ Reverse Transcriptase	1.0
Nuclease-free H_2O	4.2
Total per Reaction	**10.0**

A 2X master mix was set up according to table 2.8 with each constituent volume adjusted accordingly depending upon the number of reactions required. One extra volume of each component was added to the master mix to account for any pipetting error during reaction set up. The complete mix was gently mixed and placed on ice.

A 10 µl aliquot of the 2X master mix was dispensed into each reaction tube, to which 10 µl of the test RNA sample (100 ng/µl) was added and mixed by aspiration. This resulted in each cDNA synthesis reaction being primed with 1 µg of RNA. The reaction tubes were spun at 13,000rpm for 15 sec to collect all reaction contents. The reaction was stored on ice until loading into a thermocycler set to the conditions in table 2.9.

Table 2.9: Thermocycling conditions for cDNA synthesis

	Step 1	Step 2	Step 3	Step 4
Temperature	25°C	37°C	85°C	4°C
Time	10 min	120 min	5 min	∞

The resulting cDNA can either be immediately used for qPCR analysis or can be stored at -20 °C for a number of weeks for future use.

2.6 Quantitation of DNA and RNA

DNA and RNA were quantified using a Nanodrop spectrophotometer ND-1000. The nucleic acid concentration was measured at OD_{260} with an absorbance of 1.0 representing 50 μg/ml of DNA or 40 μg/ml of RNA. Samples were also analysed at OD_{280} to detect protein contamination and OD_{230} to detect salt/solvent contamination. The 260:280 and the 260:230 ratios will be in the region of 1.8-2.0 in a pure nucleic acid preparation. Lower ratios represent contamination with either protein or salt.

2.7 Agarose Gel Electrophoresis

Agarose gels typically prepared at a concentration of 0.7-2 % were used to analyse DNA. Gels were prepared in 1X TAE buffer, allowed to set and run in a Biorad horizontal gel apparatus using 1X TAE as the running buffer. Loading dye was mixed with the DNA samples prior to loading which facilitated accurate loading of samples and visualisation of the migration progress during electrophoresis. Electrophoresis occurred at 120 V for 20-40 min depending on the gel size and the degree of separation required. Gels were stained for 20 min in ethidium bromide solution. Visualisation was achieved with a UV transilluminator (Syngene Gene Snap) and images recorded. Gel Star gel stain was used as an alternative to ethidium bromide when it was necessary to excise DNA from the gel.

2.8 Competent Cells

Two methods for preparing competent cells were routinely employed, namely the RbCl method and the electrocompetent method. The RbCl method yields large volumes of chemically competent cells generally with an efficiency range between 10^5 to 10^8 colonies/μg of pBR322 DNA. Electrocompetent cells were prepared when transformation efficiency was of greater importance. Electrocompetent cells usually yielded efficiencies between 10^9-10^{10} colonies/μg of pBR322 DNA.

2.8.1 RbCl Competent Cells

A glycerol stock of an *E. coli* strain was streaked on LB agar and incubated overnight at 37 °C. A 10 ml aliquot of LB broth was inoculated with a single colony from this plate and incubated at 37 °C overnight in an orbital shaker at 200 rpm. A 1 ml inoculum of this overnight culture was added to 100 ml of pre-warmed SB and incubated at 37°C at 200 rpm until an OD_{600} of 0.5 was reached. The flask was immediately placed on ice for 5 min. The chilled culture was then transferred to a sterile 250 ml centrifuge tube and centrifuged in a Beckman JA-21 centrifuge at 3,000 rpm ($2500g_{av}$) at 4 °C for 5 min. The cell pellet was carefully suspended in 30 ml of ice cold TFB1 buffer, placed on ice for 90 min and centrifuged as before. The cell pellet was gently suspended in 4 ml of ice cold TFB2 buffer. The cell suspension was aliquoted into 200 µl aliquots in microfuge tubes, flash frozen in liquid nitrogen and stored at -80 °C.

2.8.2 Preparation of Electrocompetent Cells

Before this method was carried out all vessels to be used during the procedure were washed thoroughly and rinsed 3 times with deionised sterile H_2O to reduce salt contamination. A glycerol stock of an *E. coli* strain was streaked on LB agar and incubated overnight at 37 °C. A 5 ml aliquot of SB broth was inoculated with a single colony from this plate and incubated overnight at 37 °C at 200 rpm. A 1 ml inoculum from this overnight culture was used to inoculate two 250 ml aliquots of pre-warmed SB broth in 1 L flasks. These cultures were grown at 37 °C and 200 rpm until an OD_{600} of 0.75 was reached. The flasks were then cooled on ice for 15 min. The cultures were transferred to two 250 ml sterile centrifuge tubes and the cells pelleted in a Beckman JA-21 centrifuge at 3,000 rpm at 4 °C for 15 min. The pelleted cells were suspended in 250 ml of ice cold deionised sterile H_2O and centrifuged as before. The cells were then suspended in 100 ml of ice cold deionised sterile H_2O, the cell suspensions transferred into one centrifuge tube and centrifuged as before. The cell pellet was then suspended in 25 ml of ice cold 10% glycerol in deionised sH_2O and centrifuged as before. Finally the cell pellet was suspended in 1 ml of sterile 10% glycerol in deionised sH_2O and the cells aliquoted into 90 µl aliquots in sterile microfuge tubes. These were flash frozen in liquid nitrogen and stored at -80 °C. Each 90 µl aliquot is sufficient for 1-2 transformations.

2.8.3 Transformation of Chemically Competent Cells

A 1-5 µl aliquot of plasmid DNA to be transformed was added to 200 µl of competent cells that had been thawed on ice. The cells were kept on ice for 30 min and heat shocked at 42 °C for 30 sec. The cells were immediately returned to ice for 2 min. An 800 µl aliquot of SOC media was added to the cells and the mixture incubated in a 37 °C water bath for 1 hr. A 100 µl aliquot of the resulting transformation mixture was plated on the appropriate selective media and incubated at 37 °C overnight.

2.8.4 Transformation of Electrocompetent Cells

A 1 µl aliquot of a ligation mixture was added to a volume of electrocompetent cells that had been thawed on ice. An electroporation cuvette was also cooled on ice. The cells were transferred to the electroporation cuvette and all air bubbles removed by repeatedly tapping the cuvette on a hard surface. The electric pulse was delivered by a Biorad Micropulsar set to the EC2 setting. This setting delivered a pulse best suited to the *E. coli* JM109 strain as this strain requires a 0.2 cm cuvette for efficient transformation and corresponds to a 2.5 kV pulse. Immediately after electroporation, 800 µl of SOC media was added to the cuvette. This suspension was then transferred to a sterile microfuge tube and then incubated in a 37 °C water bath for 1 hr. A 100 µl aliquot was plated onto appropriate selective media and incubated at 37 °C overnight.

2.8.5 Determination of Competent Cell Efficiency

Competent cell efficiency is defined in terms of the number of colony forming units per µg of plasmid DNA used. A 25 ng/µl stock of pBR322 plasmid DNA was diluted to 250 pg/µl, 25 pg/µl and 2.5 pg/µl. A 1 µl aliquot of each dilution was transformed as described above. The cell efficiency was calculated from the number of colonies obtained, taking into account the dilution factor and the fraction of culture transferred to the spread plate.

2.8.6 Sterilisation and Cleaning of Electroportion Cuvettes

Used electroporation cuvettes were washed with copious amounts of bleach and then rinsed with sterile dH_2O six times. Each cuvette was filled with 0.25 M HCl and let stand for 2 hrs. The cuvettes were washed with sH_2O as before and boiled in dH_2O for 10 min. Immediately after removal from the bath the cuvettes were rinsed with 100% molecular grade ethanol and allowed to dry in the laminar flow cabinet. It is essential that cuvettes are dry prior to storage and that the highest grade water is used as residual salts will cause arcing during electroporation.

2.9 Storing of bacterial stocks

Glycerol stocks were prepared for each strain and stored in duplicate. A 1 ml aliquot of an exponentially growing culture was added to 0.5 ml of sterile 80% glycerol in a cryogenic tube. This was then mixed and stored immediately at -80 °C. Strains harbouring plasmids were always stored on the appropriate selective antibiotic when stocked. Working stocks were stored on agar plates at 4 °C.

2.10 Bacterial conjugation by Triparental Mating

Recipient *S. meliloti* strains were grown to late log phase in TY broth, while *E. coli* donor and helper strains were grown to late log phase in LB broth. *E. coli* donors (750 µl) were mixed with an *E. coli* strain (750 µl) harbouring the helper plasmid pRK600 (Simon et al., 1986). The mixture was pelleted at 6,000 rpm for 1 min, suspended in 100 µl of LB broth and spotted onto the centre of a LB agar plate. Following incubation overnight at 37 °C, the bacteria were suspended in 3 ml of LB broth. Then 750 µl of the recipient strain was mixed with 750 µl of the helper/donor *E. coli* strain mixture and pelleted as before. The pellet was suspended in 100 µl of TY broth and spotted onto the centre of a TY plate. Following overnight incubation the bacteria were suspended in 2 ml of TY broth and dilutions were plated onto appropriate selective media. As controls, the helper/donor *E. coli* mix and the recipient strains were carried through the entire procedure separately. The controls were then plated onto the same selective media to ensure that the selective pressure was stringent enough.

2.11 Siderophore Detection

The following two methods were routinely used to detect/measure siderophore production. Care was taken to avoid contamination of solutions with iron as excess iron greatly diminishes the sensitivity of both methods. The end colour of both procedures is a dark blue with bright blue indicating excess iron. A colour change from dark blue to orange was observed upon addition of siderophore or other strong chelator such as EDTA.

2.11.1 Chrome Azurol S (CAS) medium (Schwyn and Neilands, 1987)

CAS medium was used for the direct detection of bacterial siderophore production.

All glassware used in the preparation of CAS medium and related solutions were thoroughly washed three times with 2 M HCl and rinsed with iron free dH_2O.

- Solution 1:

PIPES	30.24 g
Sucrose	2 g
KH_2PO_4	0.03 g
Bacteriological Agar	15 g
dH_2O (Fe-)	875 ml

pH 6.8 (adjusted with 50% w/w NaOH)

- Dye Solution

60.5 mg of CAS was dissolved in 50 ml of PIPES (1 mM pH 5.6) and mixed with 10 ml of an iron(III) solution (1 mM $FeCl_3$ 10 mM HCl). While stirring, this CAS-iron solution was slowly added to 72.9 mg HDTMA dissolved in 40 ml PIPES (1 mM pH 5.6). The resultant dark blue liquid was then autoclaved.

- Deferration of Casamino Acids

Casamino acids were deferrated according to the method described by Waring (1942). The casamino acids were dissolved in 200 ml of ultra-pure water and filtered into a glass-stopped separatory funnel. 8-OH quinolone (5 mg) was dissolved in 1 ml of chloroform, poured into the funnel, shaken vigorously and allowed to stand for 5 min. Approximately 3 ml of chloroform was poured into the funnel, shaken vigorously for 1 min and then rotated for 30 sec to cause droplets of chloroform containing 8-OH quinolone-iron complexes to coalesce and settle. The chloroform layer was drawn off and the casamino acid solution washed twice with 3 ml of chloroform. The entire extraction was repeated 3 times. This procedure removed all contaminating iron.

- Completion of media

After autoclaving, solution No 1 was allowed to cool. 10 ml of the deferrated casamino acids (10%) and 10 ml of a solution containing biotin, thiamine and panthothenic acid (10 mg/L each) were added. Also added were 10 ml of 0.2 M $MgSO_4.7H_2O$ solution and 7.8 ml of $CaCl_2.H_2O$ solution (1 M). The dye solution was then added by careful addition along the glass wall with enough agitation to achieve mixing without the generation of foam.

2.11.2 Chrome Azurol S liquid assay

All solutions and dilutions were made with 1 mM PIPES pH 5.6 unless otherwise stated. A 3 ml volume of 10 mM HDTMA was placed in an iron free 50 ml volumetric flask and made up to approximately 10 ml. In a separate vessel 750 µl of ferric iron solution (1 mM $FeCl_3$, 10 mM HCl) was added to 3.75 ml of 2 mM CAS solution and the mixture slowly added to the 50 ml volumetric with mixing. In a separate vessel 2.153 g of anhydrous piperazine was partially dissolved in 10 ml of PIPES solution to which 3.125 ml of 12 M HCl was added (Caution: Highly exothermic reaction, HCl must be added dropwise to avoid flash boiling and must never be added directly to anhydrous piperazine). This solution was added to the 50 ml volumetric flask and the volume made up to 50 ml with PIPES solution. Upon the addition of the piperazine-HCl the solution will turn from a purple/wine colour to a dark blue colour.

To test a solution for the presence of siderophore, 500 µl of test solution was added to 500 µl of CAS solution. To this 10 µl of shuttle solution (0.2 M 5-sulfosalicyclic acid) was added and the mix incubated at 37 °C for 30 min to allow for complete development. The unreacted dye has a high absorbance at OD_{630} and a low absorbance at OD_{420}. The siderophore removes the iron from the dye resulting in an absorbance decrease at OD_{630} and increase at OD_{420} representing a colour change from blue to orange.

2.12 Siderophore production and analysis

2.12.1 Production of rhizobactin 1021

A *S. meliloti* strain of interest was streaked from a glycerol stock onto TY agar and incubated at 30 °C for three days. Antibiotics were included to maintain selective pressure for plasmid containing strains. A 5 ml aliquot of TY medium was inoculated with a single colony and incubated at 30 °C at 200 rpm for 2 days. A 1 ml volume of this 5 ml culture was used to inoculate 100 ml of Trizma medium containing 200 µM 2,2'dypridyl and incubated at 30 °C at 200 rpm for 2 days. The cells were removed by centrifugation at 6,000 rpm for 15 min and the supernatant transferred to a sterilised 250 ml Duran, tested for the presence of siderophore and cooled to 4 °C for short term storage.

2.12.2 Purification of rhizobactin 1021

Amberlite XAD-4 resin was used to extract siderophore from the supernatant. To prepare the resin for use it was washed in 100% ethanol at 47 °C for 30 min. This was done to remove storage salts that will reduce the resin's binding capacity. After incremental optimisation it was found that 4 g of resin per 100 ml of *S. meliloti* 2011 supernatant was sufficient for adsorption of all siderophore present. Four grams of washed XAD-4 resin was added to supernatant containing siderophore and placed at 4 °C for 24 hr. The resin was collected by vacuum filtration through a Whatman No.1 filter. The resin was washed with 50 ml of ice cold molecular grade H_2O followed by 12 ml washes with methanol at the following concentrations: 20%, 35% and 50%. All washes were carried out under a mild vacuum as the methanol washes may remove

siderophore if left for a protracted length of time. The resin was collected and placed in a 30 ml sample tube. Elution was achieved by subsequent washing with methanol/propanol/H_2O at a ratio of 30:10:1 at 47 °C. A 10 ml wash for 30 min was carried out first followed by four 6 ml washes for 30 min with the supernatant collected after each wash. The eluent, ~35 ml was dried to completion using mild heating and reduced pressure in a speed vacuum dryer. The residue was suspended in 6 ml of PBS or H_2O depending on downstream application and stored at -20 °C.

Preparative High Pressure Liquid Chromatography (HPLC) was carried out on the extracted siderophore to obtain a highly pure siderophore sample. HPLC was carried out on an Agilent 1100 Series HPLC fitted with a C-18 reverse phase column. Various concentrations of methanol/H_2O were used as the mobile phase to allow for the cleanest preparation of siderophore.

2.12.3 Analysis of siderophore by ESI-MS

Purified siderophore samples were first analysed for the presence of siderophore by the liquid CAS assay. Upon confirmation of the presence of siderophore the sample was then prepared for analysis by direct injection on a electrospray ionisation mass spectrometer.

Samples were passed through a 0.22 μm filter to ensure no large particulates were present. Particulates would result in the ESI apparatus becoming blocked and will result in instrument failure. After filtration, HPLC grade formic acid was added to a final concentration of 0.1% v/v. Formic acid is essential to this analysis as it protonates the molecules in the sample allowing for them to be detected by the mass spectrometer. Failure to include formic acid will only allow naturally occurring salts to be observed as these naturally form charged molecules.

The instrument details and settings for analysis are described in section 3.9.3 as part of the analysis of siderophore extracts.

2.13 Enzymatic Reactions

2.13.1 Enzymes and Buffers

All enzymes and their relevant buffers were obtained from the Promega Corporation, Invitrogen Life Technologies, New England Biolabs or Sigma-Aldrich Corporation and were used in accordance with the manufacturer's instructions.

2.13.2 RNase preparation

DNase free RNase was dissolved at a concentration of 10 mg/ml in 10 mM Tris-HCl (pH 7.5) and 15 mM NaCl. The solution was then dispensed into aliquots and stored at -20 °C.

2.13.3 Primer preparation

Primers were purchased from IDT DNA technologies as a lyophilised power and were suspended in TE buffer prior to use. The correct volume of TE was calculated from the amount of the primer in the tube as this varies from primer to primer depending on production yield. Primers were always prepared to a 100 nM stock and diluted 1 in 6 prior to use with the working stock being 16 nM.

2.13.4 Standard PCR Reaction Mixture

Template DNA	1 µl
Primers (16 nM)	1 µl of each
Buffer (5X)	10 µl
dNTP's (10 mM)	1 µl
Sterile dH_2O	36 µl
Polymerase (1Unit/µl	1 µl

2.13.5 Standard PCR Cycling Conditions

These cycling conditions were used for the Velocity polymerase supplied by Bioline a subsidiary of Meridian Life Sciences.

Stage 1: Step 1: 95 °C for 5 min

Stage 2: Step 1: 95 °C for 30 sec

Step 2: Annealing temperature for 10 sec

Step 3: 72 °C for 20 sec/kb to be synthesised

~35 Cycles

Stage 3: Step 1: 72 °C for 10 min

2.13.6 1 kb Plus DNA Ladder (Invitrogen)

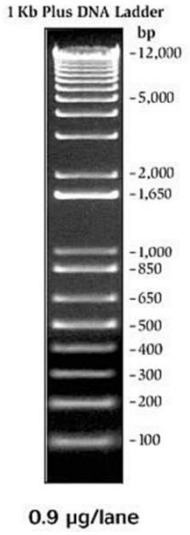


Figure 2.5: 1 kb Plus DNA Ladder

Bands of note are clearly labelled. The bands increase in 1 kb increments between 3 kb and 12 kb.

2.14 *In silico* analysis of DNA and protein sequences

- BLAST programs at NCBI http://blast.ncbi.nlm.nih.gov/Blast.cgi were used to identify homologous protein sequences.
- Promoter prediction was carried out using the Neural Network Promoter Prediction tool accessed at http://www.fruitfly.org/seq_tools/promoter.html.
- DNA and protein sequences were aligned using the Multalin program http://multalin.toulouse.inra.fr/multalin/multalin.html. More detailed alignments were obtained by MUSCLE, https://www.ebi.ac.uk/Tools/msa/muscle/ with subsequent analysis by Jalview downloaded from http://www.jalview.org/download.
- DNA sequences were analysed for restriction enzyme sites using webcutter 2.0 found at http://rna.lundberg.gu.se/cutter2/.
- Plasmid maps were drawn using the Syngene Genesnap software downloaded from http://www.syngene.com/genesnap
- Protein domains and classification analysis was carried out using the Interpro tool found at https://www.ebi.ac.uk/interpro/.
- The *S. meliloti* 1021 genome sequence is accessible from https://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime.cgi.
- Chemical structures were drawn using Chemdraw Ultra V12.0.

Chapter Three

The Role of Acetyltransferases in Rhizobactin 1021 Biosynthesis

3.1 Introduction

S. meliloti 2011 produces the siderophore rhizobactin 1021 under iron limitation. Rhizobactin 1021 is a citrate based dihydroxamate siderophore that is asymmetrically lipidated with a 2-decanoic acid moiety. The intact structure was determined by (Persmark et al. 1993) primarily through FAB-MS and various NMR methods. Figure 3.1 shows the structure of rhizobactin 1021.

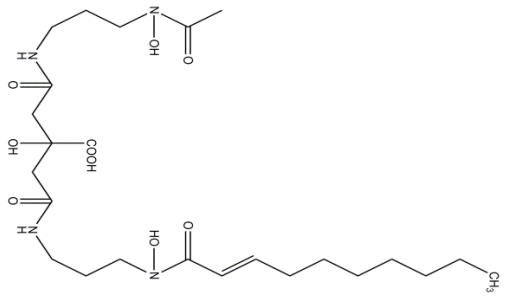


Figure 3.1: Structure of rhizobactin 1021

There have been numerous citrate dihydroxamate siderophores identified that are very similar to rhizobactin 1021. Most notable are aerobactin produced by *E. coli* K-12 and schizokinen produced by *Bacillus megaterium* and *Rhizobium leguminosarum* IARI 917. Aerobactin was the first citrate dihydroxamate siderophore to have its biosynthesis pathway experimentally determined allowing the inference of biosynthesis pathways for similar siderophores (de Lorenzo et al. 1986; de Lorenzo and Neilands 1986). Based on sequence similarity of the aerobactin biosynthesis genes to the rhizobactin 1021 genes a biosynthesis pathway for rhizobactin 1021 was proposed. Schizokinen is of interest as it is identical to the core chelating structure of rhizobactin 1021 and it is not adorned with a lipid moiety (Storey et al. 2006; Mullis, Pollack and Neilands 1971).

The biosynthesis genes of rhizobactin 1021 were mapped to a gene cluster located on the pSymA megaplasmid. Figure 3.2 is a diagrammatic representation of the rhizobactin 1021 biosynthesis and transport gene cluster. The genes *rhtX* and *rhtA* encode for the inner and outer membrane transport proteins involved in the uptake of rhizobactin 1021. The genes *rhbABCDEF* encode for the biosynthesis proteins and are transcribed as a single transcript. This transcript also includes the *rhtX* gene at the 5′ terminus. The *rhrA*

gene encodes an AraC type regulator and acts to enhance the expression of the entire cluster through DNA binding sites located upstream of *rhtX* and *rhtA* (Lynch et al. 2001; O Cuiv et al. 2004).

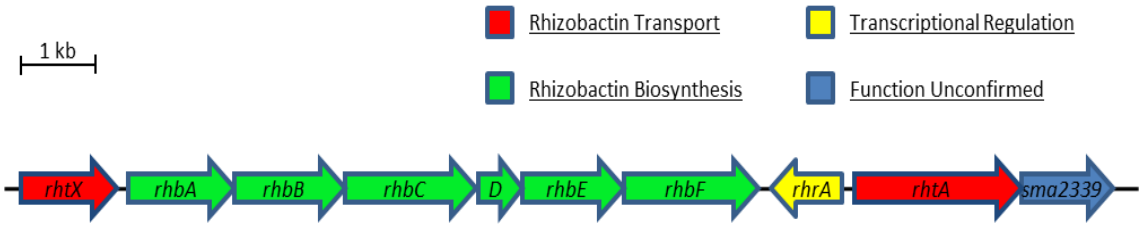
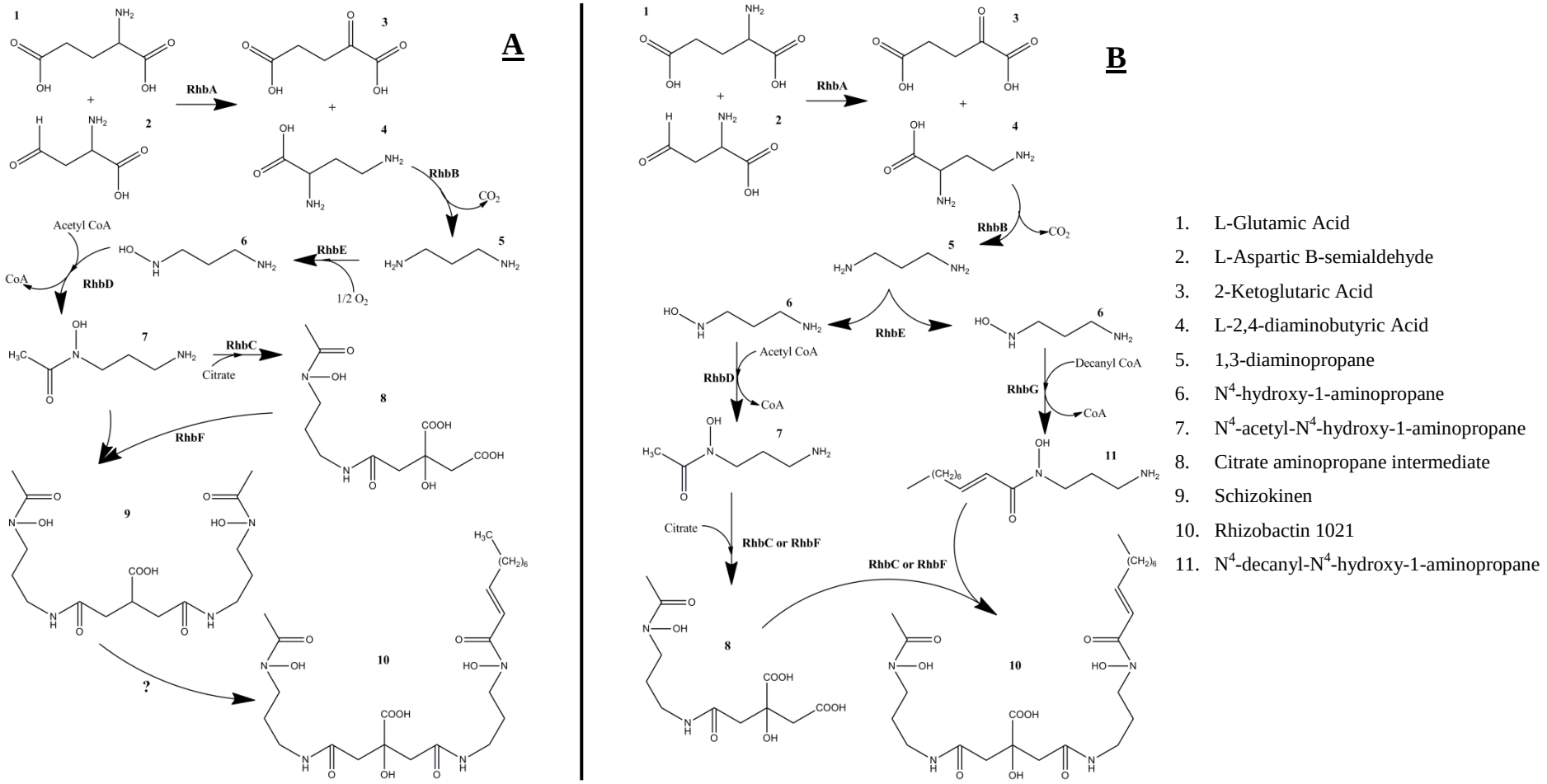


Figure 3.2: Organisation of the rhizobactin 1021 biosynthesis and transport genes

There is one uncharacterised gene, *sma2339* which forms part of the rhizobactin 1021 biosynthesis and transport region. This gene has been assigned a putative function as a siderophore biosynthesis protein based on a C-terminal acetyltransferase function. It is also likely that *sma2339* is co-transcribed with *rhtA* as there is no intergenic region between the genes. From the putative assignment of a biosynthetic function and the gene location *sma2339* has been intuitively incorporated into the rhizobactin 1021 biosynthesis pathway without direct characterisation. Lynch et al. (2001) acknowledged the presence of *sma2339* as a partial sequencing read and a transposon mutant showed no apparent change in siderophore production but it was speculated that schizokinen may have been produced in this mutant resulting in the positive CAS response.

Based on the work carried out by Lynch et al. (2001) a full biosynthesis pathway for rhizobactin 1021 was proposed. This model is inferred from the aerobactin biosynthesis pathway to which *rhbCDEF* are homologous and sequence similarity between *rhbAB* to genes responsible for 1,3-diaminopropane production in *Acinetobacter baumannii*. Notable in this pathway is that Sma2339 is not included due to no CAS deficient phenotype being observed in an *sma2339* insertion mutant and also due to the *S. meliloti* genome sequence being published after the identification of the rhizobactin regulon (Barnett et al. 2001, Capela et al. 2001, Finan et al. 2001, Galibert et al. 2001). Challis (2005) proposed a modified biosynthesis pathway based on the Lynch et al. (2001) model with the major modification being the inclusion of Sma2339, named RhbG in the role of addition of the decenoic acid moiety. The two proposed biosynthesis pathways are shown in figure 3.3 with the Lynch model labelled A and the Challis model labelled B.



1. L-Glutamic Acid
2. L-Aspartic B-semialdehyde
3. 2-Ketoglutaric Acid
4. L-2,4-diaminobutyric Acid
5. 1,3-diaminopropane
6. N^4-hydroxy-1-aminopropane
7. N^4-acetyl-N^4-hydroxy-1-aminopropane
8. Citrate aminopropane intermediate
9. Schizokinen
10. Rhizobactin 1021
11. N^4-decanyl-N^4-hydroxy-1-aminopropane

Figure 3.3: Proposed biosynthesis pathways for rhizobactin 1021

The two pathways are identical for the first number of steps. The RhbA protein catalyses the production of L-2,4-diaminopropane from the precursors L-glutamic acid and L-aspartic B-semialdehyde. The by-product 2-ketoglutaric acid is not used in the biosynthesis pathway. The L-2,4-diaminopropane is decarboxylated by RhbB to produce 1,3-diaminopropane which is in turn converted to N^4-hydroxy-1-aminopropane by the oxygenase activity of RhbE. It is at this point the pathways diverge.

Lynch Model

The N^4-hydroxy-1-aminopropane is converted to N^4-acetyl-N^4-hydroxy-1-aminopropane via an acetyltransferase activity provided by RhbD using acetyl CoA as a substrate. This is then added to a single citrate molecule by RhbC to create a citrate aminopropane intermediate. The RhbC product is converted to schizokinen by the action of RhbF with the addition of a second N^4-acetyl-N^4-hydroxy-1-aminopropane. The final step is left open to speculation with respect to the mechanism for addition of the lipid.

Challis Model

Note that RhbE apparently producing more than one N^4-hydroxy-1-aminopropane is only for clarity in drawing the model. Similar to the Lynch model N^4-hydroxy-1-aminopropane is acetylated by RhbD. N^4-hydroxy-1-aminopropane is also acetylated by RhbG using a decanyl CoA to produce N^4-decanyl-N^4-hydroxy-1-aminopropane. The exact roles of rhbC and rhbF are left uncertain with it being proposed that they act to assemble the products of RhbD and RhbG into the final rhizobactin 1021 structure.

As mentioned above, the major divergence in the models is the assignment of Sma2339 as a key biosynthesis protein. The uncertainty of the roles of RhbC and RhbF is due to their sequence similarity to one another however basing their activity strictly on their homologs in the aerobactin synthesis pathway, IucA and IucC respectively, then the Lynch model is most likely correct, in that RhbC acts first forming the amide bond between N^4-acetyl-N^4-hydroxy-1-aminopropane and citrate. The RhbF protein then recognises this modified citrate structure and catalyses the formation of a second amide bond thus completing the core siderophore structure. It has not been concluded whether RhbC and RhbF, or indeed IucA and IucC, act sequentially to one another as shown above or if they act in unison (de Lorenzo and Neilands 1986).

Resulting from the uncertainty surrounding the role of Sma2339 in rhizobactin 1021 synthesis it was decided to experimentally characterise the phenotype of a *sma2339* knockout mutant in respect to rhizobactin 1021 production. The role of RhbD in the biosynthesis of rhizobactin 1021 was also assessed to ensure that the results obtained from the *sma2339* knockout analysis were not masked by the RhbD protein compensating. RhbD and Sma2339 are both classed as acetyltransferases by bioinformatics and they may be able to complement each other's activity.

3.2 Comparison of the rhizobactin 1021 biosynthesis operon to similar siderophore operons

As stated earlier the rhizobactin 1021 biosynthesis pathway from Lynch et al. (2001) is inferred from the aerobactin biosynthesis pathway based on the sequence similarities of the synthesis proteins. To analyse the level of conservation of this biosynthesis pathway, bacterial strains that have been shown to produce citrate dihydroxamate siderophores and have a closely related strain sequenced were analysed for the presence of a synthesis operon similar to that found in figure 3.2. The selected siderophore structures are given in figure 3.4 and table 3.1.

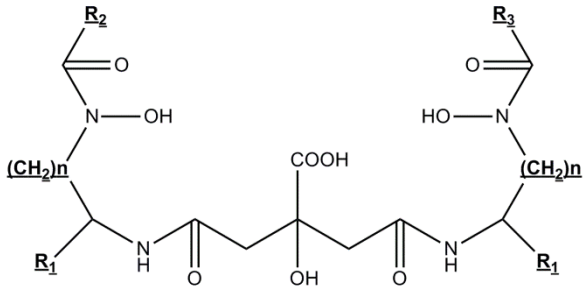


Figure 3.4: Core structure of citrate dihydroxamate siderophores

Table 3.1: Variable features of citrate dihydroxamate siderophores

Siderophore	R_1	R_2	R_3	n
Rhizobactin 1021	H	CH_3	(*E*)-2-decenoic acid	2
Schizokinen	H	CH_3	CH3	2
Acinetoferrin	H	(*E*)-2-octenoic acid	(*E*)-2-octenoic acid	2
Synechobactin A	H	CH_3	dodecanoic acid	2
Synechobactin B	H	CH_3	decanoic acid	2
Synechobactin C	H	CH_3	octanoic acid	2
Aerobactin	COOH	CH_3	CH_3	4
Ochrobactin A	COOH	(*E*)-2-octenoic acid	(*E*)-2-decenoic acid	4
Ochrobactin B	COOH	octanoic acid	(*E*)-2-decenoic acid	4
Ochrobactin C	COOH	(*E*)-2-decenoic acid	(*E*)-2-decenoic acid	4

*Sources of siderophore structures are discussed below.

The main difference between the core chelating structures of the siderophores is the length of the dihydroxamate appendages. This is represented by the n value with an n value of 2 most likely representing siderophores constructed from 1,3 diaminopropane, and those with a value of 4 most likely constructed from L-lysine. Schizokinen and aerobactin would represent the base for all the other siderophores in table 3.1 with the other morphologies being acetyl chain length changes at positions R_2 and R_3. In essence

rhizobactin 1021, acinetoferrin and synechobactin A-C are based on a schizokinen core structure with ochrobactin A-C based on an aerobactin core structure.

Table 3.1 is not an exhaustive list of citrate dihydroxamate based siderophores but represents a selection that has the producing strain or a closely related strain sequenced. This was done to ensure a level of confidence in that the operons studied were likely to encode for siderophores of known structure. A general BLASTp analysis would have resulted in the identification of homologs for which the siderophore structure was not confirmed which may have skewed the analysis. Table 3.2 outlines the producing strain of each siderophore and the source of the gene sequence used for identifying the biosynthesis operons.

Table 3.2: Siderophore producing strains and the corresponding genome sequence source

Siderophore	Producing Strain	Sequence Source
Rhizobactin 1021	*S. meliloti* 1021	*S. meliloti* 1021
Schizokinen	*B. megaterium* ATCC 19213	*B. megaterium* DSM 319
Acinetoferrin	*A. haemolyticus* ATCC 17906	*A. haemolyticus* ATCC 19194
Synechobactin A-C	*Synechococcus* Sp. PCC 7002	*Synechococcus* Sp. PCC 7002
Aerobactin	*E. coli* ColV-K30	Plasmid ColV-K30
Ochrobactin A-C	*Ochrobactrum sp* SP18	*Ochrobactrum anthropi*

Of the siderophores presented in table 3.1 only three have had the biosynthesis genes physically characterised by mutagenesis. Rhizobactin 1021 was structurally characterised by Persmark et al. (1993) with the biosynthesis operon identified by Lynch et al. (2001). Aerobactin was originally identified as a product from *E. aerogenes* and characterised by Gibson F. (1969). Subsequent to its initial discovery aerobactin was found to be produced by *E. coli* harbouring the pColV-K30 plasmid and it was from the pColV-K30 plasmid that the biosynthesis cluster was identified (de Lorenzo et al. 1986; de Lorenzo and Neilands 1986). The structure of acinetoferrin was determined by Okujo et al. (1994) and the biosynthesis cluster subsequently identified by Funahashi et al. (2013). The remaining siderophores schizokinen, synechobactin A-C and ochrobactin A-C have only been structurally characterised without the biosynthesis genes experimentally determined (Martin et al. 2006; Storey et al. 2006; Mullis, Pollack and Neilands 1971; Ito and Butler 2005). For strains that did not have a sequence available a related strain was used for analysis. The biosynthesis genes for acinetoferrin were identified in *A. haemolyticus* ATCC 17906 but the strain *A. haemolyticus* ATCC

19194 was used for the genome sequence. This was done as only the siderophore biosynthesis genes were sequenced in *A. haemolyticus* ATCC 17906 whereas the entire genome sequence was available for *A. haemolyticus* ATCC 19194. BLASTp searching showed that the entire biosynthesis pathway was intact in *A. haemolyticus* ATCC 19194 and it was decided to use this strain to ensure that possible significant genes located outside of the biosynthesis operon were not overlooked.

3.2.1 Comparison of the gene arrangement of biosynthesis clusters

Using the protein sequences of RhbABCDEF and Sma2339, BLASTp analysis was carried out on the sequences of the strains in table 3.2. Priority was given to results that linked to proteins encoded from apparent operons or gene clusters. The transport proteins were not essential to this analysis but are included where they were found as part of an operon or close to the operon. A diagrammatic gene map for the identified biosynthesis gene clusters for each strain is given in figure 3.5.

The biosynthesis operons are arranged into two groups based on whether their produced siderophore is based on a 1,3 diaminopropane or L-lysine hydroxamate source. As expected, each of the operons includes homologs to the *rhbCEF* genes as these are responsible for the assembly of the siderophore core structure. Only the strains producing 1,3 diaminopropane based siderophores were found to have homologs of *rhbAB*. The *rhbAB* homologs were found as part of the main biosynthesis operon for rhizobactin 1021, schizokinen and synechobactin A-C but are located elsewhere on the genome in *A. haemolyticus* ATCC 19194. The exact positioning of these two regions in relation to one another in *A. haemolyticus* ATCC 19194 could not be determined as the genome has only been shotgun sequenced and is not yet assembled. The two genes *dat* and *ddc* are not mentioned as part of the acinetoferrin biosynthesis operon described by (Funahashi et al. 2013) leaving their role in siderophore synthesis open to speculation.

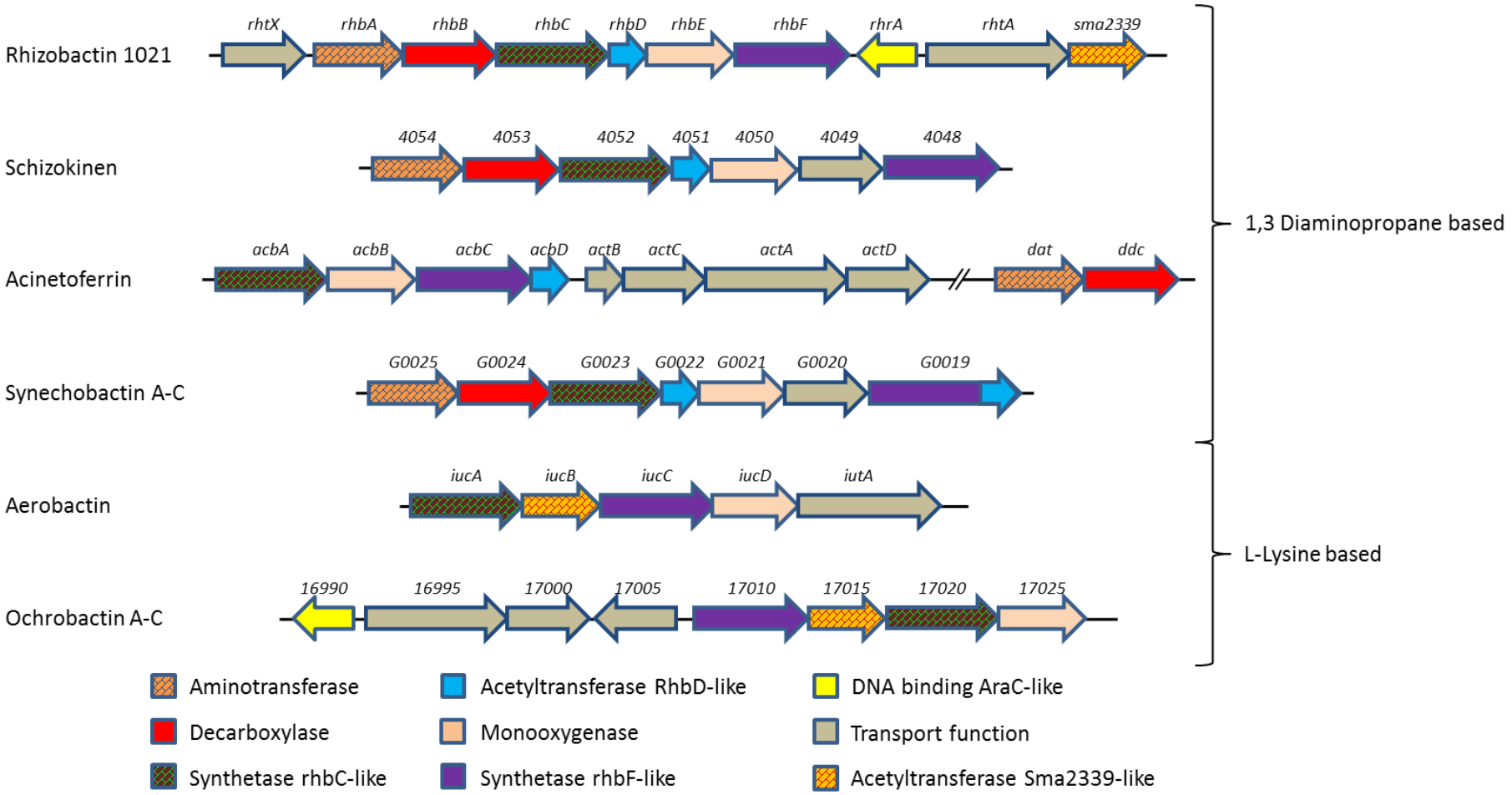


Figure 3.5: Biosynthesis operons of citrate dihydroxamate siderophores

A. haemolyticus ATCC 17906 has been shown to produce 1,3 diaminopropane irrespective of iron stress and perhaps this allows the production of acinetoferrin without the genes being present in the biosynthesis operon (Ikai and Yamamoto 1997; Yamamato et al. 1995). The acetyltransferase gene of the L-lysine based siderophore biosynthesis operons appears to be more closely related to *sma2339* than to *rhbD*. The significance of this is uncertain as there does not appear to be an obvious link between the acetyltransferase present and the acetyl chain length at position R_2 and R_3 when examining table 3.1 and figure 3.5. The absence of a clear link between acetyltransferases and acetyl chain length suggests that their role in biosynthesis may not be as clear as previously thought.

It is worth noting that *rhbD* and *sma2339* are similar to one another. The full details of this similarity will be discussed below but in brief Sma2339 has a large domain at the N-terminus that is completely missing from RhbD and its homologs. The C-terminal domain of Sma2339 is highly similar to RhbD as it is the acetyltransferase domain. The genes *iucB* and 17015 are similar in length to *sma2339* and this may be the reason for the increased similarity to Sma2339 over RhbD.

Although not strictly a part of this analysis the presence of transport proteins in each cluster is of interest. At least one major facilitator superfamily (MFS) transporter is present in each cluster except for that for aerobactin. Some of these MFS transporters are homologs to RhtX where others show very poor homology. There is also a TonB-dependent outer membrane receptor in each cluster that is from a Gram-negative source, except in *Synechococcus* SP. PCC 7002 were it appears to be located distal to the biosynthesis cluster. These transport genes will be included in further bioinformatic analysis but will not be investigated in vivo.

3.2.2 Protein homology analysis of the rhizobactin 1021 operon to related siderophore operons

Having identified biosynthesis operons for each of the siderophores, a more in depth analysis was then carried out. Each protein from the rhizobactin 1021 operon was analysed and then compared to its counterpart in the other clusters.

RhtX transporter homologs

The protein RhtX has a predicted molecular weight of 45.5 kDa and a pI of 9.84. It is the defining member of the RhtX/FptX sub-family of the Major Facilitator Superfamily. Structural analysis reveals 12 transmembrane alpha-helices that arrange themselves in the inner membrane to form a pore structure (O Cuiv et al. 2004). The RhtX protein was used as a probe and using BLASTp each of the target strains were analysed for the presence of similar transporters. Focus was given to those located close to known or putative siderophore biosynthesis operons. The data from those analyses are given in table 3.3.

Table 3.3: Proteins showing significant homology to RhtX

Target Strain	Protein	% Coverage	% Identity	e-value
***B. megaterium* DSM 319**	None	N/A	N/A	N/A
***A. haemolyticus* ATCC 19194**	ActC	90%	34%	1e-62
***Synechococcus Sp*. PCC 7002**	None	N/A	N/A	N/A
pColV-K30	None	N/A	N/A	N/A
Ochrobactrum anthropi	BG46_17000	89%	53%	1e-122

Only proteins that showed a significant match to RhtX were considered. Of the five sequences analysed only two had MFS transporters that were similar to RhtX and located close to the siderophore biosynthesis cluster. The positive hits were observed in *A. haemolyticus* and *O. anthropi* which are Gram-negative bacteria whereas *B. megaterium* is Gram-positive and *Synchococcus Sp.* is a member of the cyanobacteria with a Gram-negative cell structure. The absence of an RhtX homolog in *B. megaterium* is not surprising given that it is Gram-positive and its transporters would require different characteristics. The pColV-K30 plasmid was found to encode for aerobactin biosynthesis and outer membrane transport only, with inner membrane transport proteins encoded for on the chromosome of the harbouring strain.

The ActC protein from *A. haemolyticus* was shown to be responsible for acinetoferrin transport across the inner membrane by Funahashi et al. (2013) and BG46_17000 is likely to carry out this role in *O. anthropi*. In addition to the RhtX homologs a number of MFS transporters non-homologous to RhtX were observed by sequence gazing. ActD in *A. haemolyticus* plays a role in acinetoferrin secretion as determined by Funahashi et al. (2013) and shows homology to BG46_17005 in *O. anthropi* with 40% coverage, 22% identity and an e value of 0.022 suggesting that BG46_17005 may play a role in siderophore secretion. The two remaining MFS transporters are BMD_4049 from *B. megaterium* and G0020 of *Synechococcus Sp.* with both showing no homology to either RhtX or ActD. They show significant homology to one another with 79% coverage, 29% identity and an e value of 3e-38 indicating a highly significant match. They may represent examples of siderophore import/export proteins due to their distinct amino acid sequences and location.

RhbA and RhbB, 1,3 diaminopropane production homologs

As both the RhbA and RhbB proteins are required to produce 1,3 diaminopropane homologs were only considered to be fulfilling this role if they were located proximally to one another. The protein RhbA has a predicted molecular weight of 50 kDa and a pI of 6.53 with the RhbB protein having a molecular weight of 52.5 kDa and a pI of 6.01. The BLASTp results obtained from analysing each sequenced strain in table 3.2 for the presence of RhbA and RhbB are given in tables 3.4 and table 3.5.

Table 3.4: Proteins showing significant homology to RhbA

Target Strain	Protein	% Coverage	% Identity	e-value
***B. megaterium* DSM 319**	BMD_4054	93%	55%	8e-178
***A. haemolyticus* ATCC 19194**	Dat	93%	53%	3e-168
***Synechococcus* Sp. PCC 7002**	G0025	92%	59%	0.0
pColV-K30	None	N/A	N/A	N/A
Ochrobactrum anthropi	None	N/A	N/A	N/A

Table 3.5: Proteins showing significant homology to RhbB

Target Strain	Protein	% Coverage	% Identity	e-value
***B. megaterium* DSM 319**	BMD_4053	95%	38%	4e-113
***A. haemolyticus* ATCC 19194**	Ddc	93%	38%	5e-100
***Synechococcus Sp*. PCC 7002**	G0024	85%	40%	7e-107
pColV-K30	None	N/A	N/A	N/A
Ochrobactrum anthropi	None	N/A	N/A	N/A

As expected there are very significant matches found in each of the strains producing 1,3 diaminopropane based siderophores. There were multiple hits for RhbA homologs in each organism however in *O. anthropi* none of the corresponding genes were co-located with an *rhbB* gene or with a gene encoding a similar function and only one was co-located in the 1,3 diaminopropane siderophore producing strains. No homologs to either protein were found to be encoded on the pColV-K30 plasmid and this along with the result in *O. anthropi* correlate with their L-lysine based siderophore structure.

RhbC and RhbF, siderophore synthetase homologs

As shown before in figure 3.3 RhbC and RhbF act to assemble the hydroxamate moieties onto the citrate core of the siderophore. The RhbC protein has a predicted molecular weight of 64.6 kDa and a pI of 5.95 with the RhbF protein having a molecular weight of 67 kDa and a pI of 5.84. The homologies of the synthetases in each operon to RhbC and RhbF are given in table 3.6 and table 3.7 respectively.

Table 3.6: Proteins showing significant homology to RhbC

Target Strain	Protein	% Coverage	% Identity	e-value
***B. megaterium* DSM 319**	BMD_4052	98%	35%	5e-110
	BMD_4048	56%	21%	3e-08
***A. haemolyticus* ATCC 19194**	AcbC	91%	34%	3e-94
	AcbC	52%	21%	1e-07
***Synechococcus Sp*. PCC 7002**	G0023	97%	35%	2e-111
	G0019	29%	21%	0.01
pColV-K30	IucA	90%	28%	4e-59
	IucC	77%	25%	4e-26
Ochrobactrum anthropi	BG46_17020	96%	31%	1e-80
	BG46_17010	58%	27%	1e-15

Table 3.7: Proteins showing significant homology to RhbF

Target Strain	Protein	% Coverage	% Identity	e-value
***B. megaterium* DSM 319**	BMD_4048	96%	40%	2e-152
	BMD_4052	58%	21%	2e-07
***A. haemolyticus* ATCC 19194**	AcbC	96%	37%	3e-134
	AcbA	60%	24%	2e-13
***Synechococcus Sp*. PCC 7002**	G0019	98%	42%	4e-175
	G0023	38%	24%	4e-09
pColV-K30	IucC	95%	35%	6e-110
	IucA	42%	24%	2e-07
Ochrobactrum anthropi	BG46_17010	97%	39%	1e-141
	BG46_17020	13%	33%	8e-08

Due to the similar activity of the RhbC and RhbF proteins their homologs overlap extensively. However even with this similarity there are still distinct homologs for each of the synthetases. Whereas it is not unexpected to find close homologs in each of the biosynthesis clusters, it is surprising that there isn't a lower level of homology between RhbC and RhbF to their direct homologs in *O. anthropi* and on the pColV-K30 plasmid as these proteins recognise a L-lysine derived substrate with RhbC and RhbF recognising a 1,3 diaminopropane derived substrate. Perhaps these synthetases act

indiscriminately of the size of their substrate allowing them to assemble siderophores with varying length hydroxamate structures acetylated with variable length acetyl groups.

RhbE, monooxygenase homologs

The RhbE protein is responsible for the addition of the hydroxyl group to the amide group of 1,3 diaminopropane. The oxygen from this hydroxyl group is involved in the coordination of Fe^{3+} along with the hydroxyl group of the terminal acetyl group and the central carboxyl group provided by the citrate molecule as shown by studies on schizokinen, aerobactin and acinetobactin (Fadeev, Luo and Groves 2004; Harris, Carrano and Raymond 1979). The RhbE protein has a molecular weight of 50.5 kDa and a pI of 6.63. The details of the homologs identified by BLASTp are given in table 3.8.

Table 3.8: Proteins showing significant homology to RhbE

Target Strain	Protein	% Coverage	% Identity	e-value
***B. megaterium* DSM 319**	BMD_4050	94%	43%	2e-123
***A. haemolyticus* ATCC 19194**	AcbB	97%	41%	9e-108
***Synechococcus Sp*. PCC 7002**	G0021	94%	44%	3e-119
pColV-K30	IucD	91%	33%	2e-73
Ochrobactrum anthropi	BG46_17025	92%	37%	3e-76

As expected, due to the essential function of the hydroxyl group attached by RhbE, each of the biosynthesis operons has a clear homolog. There is a slight reduction in % coverage and % identity observed in both the pColV-K30 plasmid and in *O. anthropi* which may be attributed to its activity on L-lysine derived substrate rather than 1,3 diaminopropane.

RhrA, transcription regulator and RhtA, outer membrane receptor homologs

The RhrA protein is an AraC-like DNA binding protein and acts as a positive transcriptional regulator of the rhizobactin 1021 operon. It has a molecular weight of 35 kDa and a pI of 8.95. BLASTp analysis of the target genome sequences only identified one homolog near the biosynthesis cluster which is present in *O. anthropi*. This is BG46_16990 and shows a % identity of 33% with coverage over 98% of the protein and an e-value of 1e-53. Given the similarity and location of BG46_16990 to RhrA it is most likely a positive regulator of ochrobactin A-C production and perhaps also the outer membrane receptor. As none of the other clusters possessed a similar protein it is likely they rely solely on derepression of the cluster by iron stress mediated through a Fur like global regulator.

The outer membrane receptor for rhizobactin 1021, RhtA has a molecular weight of 80.5 kDa and a pI of 4.52. As expected homologs were found in each of the Gram-negative organisms and on the pColV-K30 plasmid as only this class of outer membrane transporter is known to transport siderophores. The results of the BLASTp analysis are given in table 3.9.

Table 3.9: Proteins showing significant homology to RhtA

Target Strain	Protein	% Coverage	% Identity	e-value
***B. megaterium* DSM 319**	None	N/A	N/A	N/A
***A. haemolyticus* ATCC 19194**	ActA	95%	33%	1e-111
***Synechococcus Sp*. PCC 7002**	G0138*	93%	30%	4e-77
pColV-K30	IutA	96%	42%	0.0
Ochrobactrum anthropi	BG46_16995	96%	26%	2e-48

*Not located as part of the synechobactin A-C cluster

The homologs to RhtA in *A. haemolyticus* and on the pColV-K30 plasmid have been characterised and are known to transport acinetoferrin and aerobactin respectively. The 26% identity for BG46_16995 to RhtA indicates a similar function which may not be identical. This may be due to the variety of possible siderophore structures that BG46_16995 is required to transport which will require structural changes in the receptor.

The result from *Synechoccus Sp*. is notable as the protein G0138 is transcribed distally to the biosynthesis operon. There are six TonB-dependent outer membrane receptors on

the same plasmid as the synechobactin biosynthesis operon in *Synechococcus Sp*. with G0128 showing the best homology. As no location based assumptions can be made about each of these receptors the transport substrate can only be determined by experimentation. This is due to BLASTp results on outer membrane receptors often being misleading as they can also show high levels of similarity to other receptors that do not transport the same or similar structured siderophores. The TonB-dependent outer membrane receptor BG46_16995 is a more certain assumption as it is located proximal to the ochrobactin biosynthesis cluster. However the only way to determine the exact substrates is by mutagenesis as there may be more transporters for each of the ochrobactins A, B and C.

RhbD and Sma2339, acetyltransferase homologs

The RhbD protein has been assigned a "capping" function in that it acetylates the N^4-hydroxy-1-aminopropane molecule created by RhbE. This activity results in the completion of the substrate recognised by RhbC which is bonded to the citrate structure. The acetylation also results in the additions of an oxygen molecule, to which iron is coordinated, to the structure of the siderophore. The Sma2339 protein has been proposed to carry out the same function as RhbD but acetylating with a decanyl moiety instead of an acetyl moiety (Challis 2005). The molecular weight and pI values for RhbD are 23 kDa and 6.84 respectively and 40 kDa and 9.04 respectively for Sma2339.

Table 3.10: Proteins showing significant homology to RhbD

Target Strain	Protein	% Coverage	% Identity	e-value
***B. megaterium* DSM 319**	BMD_4051	86%	35%	7e-38
***A. haemolyticus* ATCC 19194**	AcbD	86%	37%	4e-36
***Synechococcus Sp*. PCC 7002**	G0022	91%	49%	3e-59
	G0019	93%	27%	1e-20
pColV-K30	IucB	94%	37%	1e-40
Ochrobactrum anthropi	BG46_17015	95%	40%	5e-47

Table 3.11: Proteins showing significant homology to Sma2339

Target Strain	Protein	% Coverage	% Identity	e-value
***B. megaterium* DSM 319**	BMD_4051	43%	28%	5e-18
***A. haemolyticus* ATCC 19194**	AcbD	42%	27%	9e-16
***Synechococcus Sp*. PCC 7002**	G0022	50%	34%	2e-29
	G0019	42%	25%	6e-08
pColV-K30	IucB	74%	36%	1e-52
Ochrobactrum anthropi	BG46_17015	86%	33%	1e-36

As expected there are homologs of RhbD in each biosynthesis cluster thus confirming its central role in biosynthesis. The result in *Synechococcus Sp*. is unusual as the RhbF homolog G0019 has a distinct acetyltransferase domain at its C-terminal. Perhaps this is a variation of the rhizobactin 1021 biosynthesis cluster in that there are two acetyltransferase proteins present. The Sma2339 protein shows homology to the same proteins as RhbD however there are lower levels of sequence coverage and identity between most of the proteins. The lower % coverage for BMD_4051, AcbD, G0021 and G0019 is most likely due to the size difference between these proteins and Sma2339. Sma2339 is the largest of the acetyltransferase proteins at 370 aa with the others,

including RhbD being ~200 aa. To ensure that the greater size of Sma2339 was not skewing the BLASTp results, all of the acetyltransferase proteins were BLASTp searched against the *S. meliloti* 1021 genome. Only the C-terminus of G0019 was used in this search as the majority of the protein is homologous to RhbF and may interfere with the analysis. The results of this BLASTp search are given in table 3.12.

Table 3.12: Acetyltranferase proteins BLASTp searched against *S. meliloti*

Subject Protein	Homolog	% Coverage	%Identity	E-Value
BMD_4051	RhbD	85%	35%	7e-40
	Sma2339	81%	28%	5e-22
AcbD	RhbD	79%	37%	2e-36
	Sma2339	74%	27%	3e-16
G0022	RhbD	89%	49%	2e-59
	Sma2339	92%	34%	5e-30
G0019	RhbD	76%	27%	7e-22
	Sma2339	66%	25%	2e-08
IucB	RhbD	58%	37%	3e-37
	Sma2339	84%	36%	3e-49
BG46_17015	RhbD	52%	40%	2e-47
	Sma2339	88%	33%	7e-43
RhbD	RhbD	100%	100%	1e-145
	Sma2339	93%	37%	4e-32
Sma2339	RhbD	49%	37%	9e-32
	Sma2339	100%	100%	0.0

The homologies of the acetyltransferase proteins indicate that the majority are more similar to RhbD. However IucB shows a greater degree of similarity to Sma2339 than to RhbD and BG46_17015 displays nearly the same match significance to both RhbD and Sma2339. The results from these analyses don't show any correlation between similarity to RhbD or Sma2339 and whether their siderophore is acetylated with a fatty acid or a simple acetyl group.

This is unexpected as the characterised biosynthesis pathway for aerobactin indicates that IucB is responsible for addition of the acetyl cap on the siderophore. The addition of a lipid in place of the acetyl group would require identification of a larger lipid chain substrate over a short acetyl CoA substrate but this does not appear to result in a noticeable change in the protein. To assess whether there are any conserved features

between the proteins they were aligned by the Multiple Sequence Comparison by Log-Expectation (MUSCLE) algorithm. Figure 3.6 shows the result of the alignment using the same regions as used for the BLASTp search in table 3.12.

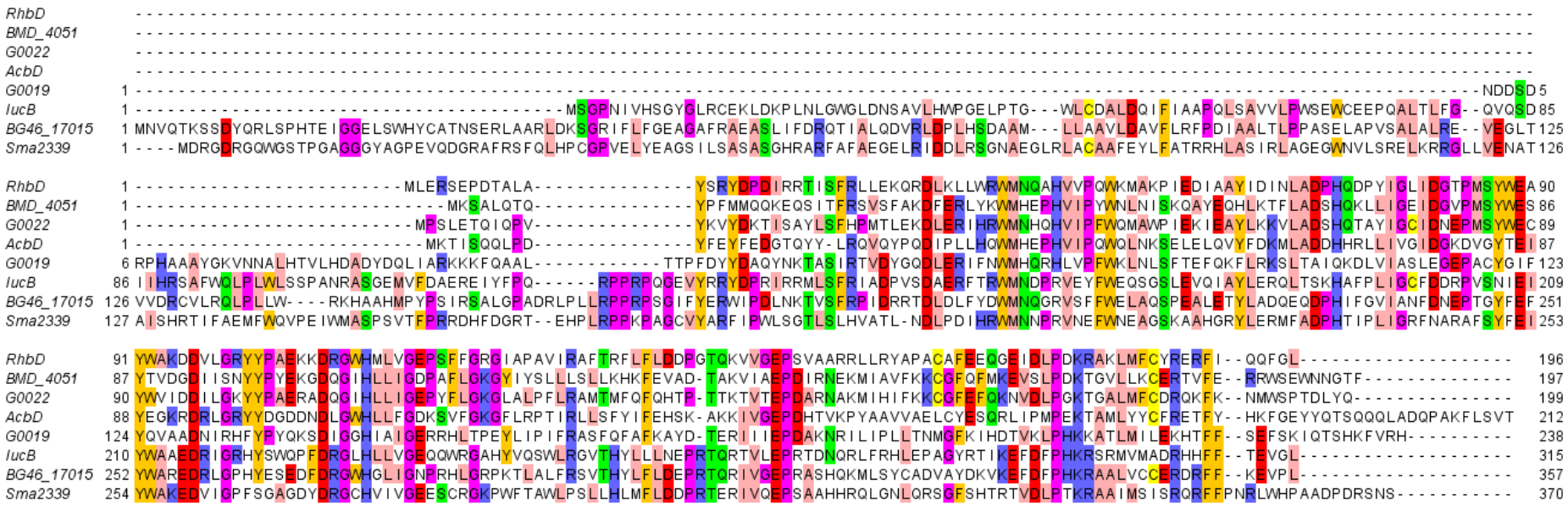


Figure 3.6: Alignment of the siderophore biosynthesis acetyltransferase proteins

The protein residues are coloured according to the Zappo colour scheme provided through the Jalview multiple sequence alignment analysis tool and editor. Only residues that displayed over 25% identity were coloured to demonstrate the level of conservation at those particular position and also to show the physical properties at highly conserved sites. Table 3.13 is a description of the Zappo colour scheme.

Table 3.13: Zappo colour scheme for multiple alignments

Residue Property	Amino Acid	Colour
Aliphatic/hydrophobic	ILVAM	Pink
Aromatic	FWY	Orange
Positive	KRH	Blue
Negative	DE	Red
Hydrophilic	STNQ	Green
Conformationally Special	PG	Purple
Cysteine	C	Yellow

There do not appear to be any regions conserved in the RhbD protein homologs and that are missing from the Sma2339 protein homologs. There is no feature identified that would allow speculation of any functional bias to one protein type or the other.

3.2.2.1 Direct comparison of RhbD to Sma2339

As mentioned in section 3.2.1 RhbD and Sma2339 are homologous to one another with the C-terminal of Sma2339 being homologous to the RhbD protein. This is due to the size differences between the two proteins as RhbD is 196 amino acids in length and Sma2339 is 370 amino acids in length. To demonstrate the similarities between the two proteins a protein alignment was carried out using the MUSCLE online tool. The results of this alignment are given in figure 3.7.

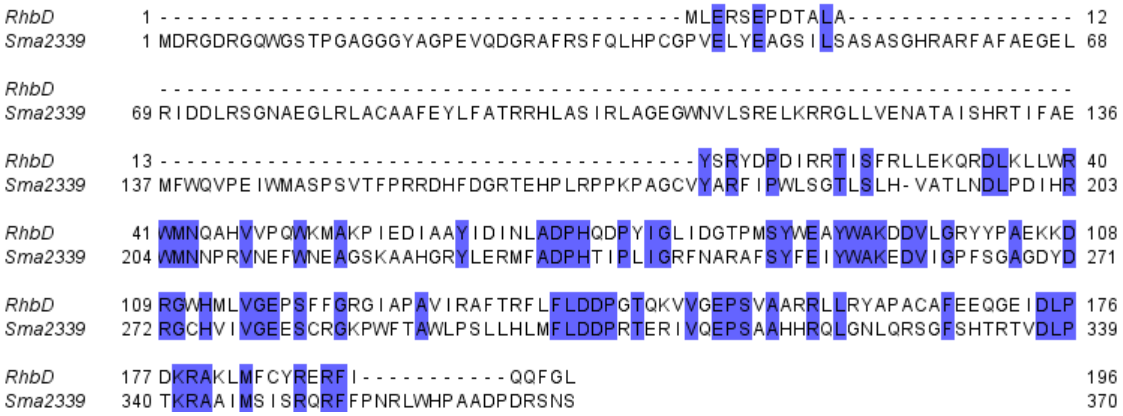


Figure 3.7: Alignment of RhbD and Sma2339 protein sequences

The conserved residues between the two proteins are highlighted in blue. As can be observed there are a large number of conserved sites at the C-terminus of Sma2339 due to the predicted acetyltransferase function attributed to this region.

3.2.3 Summary of bioinformatic analysis of siderophore biosynthesis proteins

The comparison of the rhizobactin 1021 biosynthesis cluster to those of similar structured siderophores identified key homologs in each cluster. It was determined that each of the citrate 1,3 diaminopropane based siderophores possesses homologs to RhbA and RhbB which are responsible for 1,3 diaminopropane production. The strains producing L-lysine based siderophores also have homologs to these genes but they are located distally to one another indicating that they are unlikely to interact with one another. Supplementary to 1,3 diaminopropane production it was observed that each of the clusters had two siderophore synthetases homologous to RhbC and RhbF and an oxygenase protein homologous to RhbE with the functions of each described in figure 3.3.

Based on the assigned role of acetyltransferases in siderophore biosynthesis it was not unexpected to find homologs in each cluster. However according to the known role of acetyltransferase activity in siderophore synthesis shown in figure 3.3, it would be expected that there would be changes in the protein depending on the length of the acetyl chain to be added. No correlation between acetyl chain length and acetyltransferase sequence was observed.

As a clear difference between the acetyltransferase structure and the acetyl chain length could not be identified the only way to determine the actual function of RhbD and Sma2339 was through mutagenesis. A strategy was designed to knockout *sma2339* and *rhbD* separately and to create a double knockout mutant of the two genes. Extracts were then prepared from each of these mutants and assessed for the presence of siderophore and then subjected to analysis by ESI-MS. Extracts from wild type and a siderophore deficient strain, *S. meliloti* 2011*rhbA*, were used as controls to allow rhizobactin 1021 to be identified in each sample.

3.3 The principles of triparental mating and mutagenesis of *S. meliloti*

The availability of sequenced microbial genomes and the advances of PCR derived techniques have allowed for the direct targeting of regions for mutagenesis. Transposable elements such as Tn5*mob* and Tn10*mob,* and antibiotic resistance cassettes are now widely used in creating gene knockouts through insertional inactivation. These advances along with the isolation of a variety of genes responsible for antibiotic resistance have allowed the creation of multiple site directed mutations in the same organism. Suicide vectors have been developed that lack broad host range functions of plasmids which facilitates the selection of recombinant events at a detectable frequency. One such vector, pJQ200sk has been developed that is compatible with carrying out mutagenesis in rhizobial strains.

For mutagenesis in *S. meliloti* the pJQ200sk vector is first modified in an *E. coli* host and then conjugated into *S. meliloti* by manipulation of the mobilisable functions of pJQ200sk. The functions that allow for conjugation are expressed *in trans* from the narrow host range plasmid pRK600 and are referred to as the *tra* genes. The pRK600 plasmid is referred to as the helper plasmid. A schematic of the process of triparental mating and the subsequent recombination events that result in mutagenesis is presented in figures 3.8 and figure 3.9. The pJQ200sk vector integrates into the chromosome as it cannot replicate in the *S. meliloti* host. A 2 kb region is sufficient to allow the pJQ200sk vector to integrate into the genome by homologous recombination with the hosts DNA. When an antibiotic resistance cassette is to be incorporated into the genome it is placed centrally in this 2 kb region with the 1 kb on either side being sufficient to allow recombination. The direction of the antibiotic cassette should be noted, as the direction of transcription from the resistance gene can cause or mask polar effects. If the cassette is in the same orientation as the mutated gene, polar effects can be masked as the cassette promoter may transcribe the downstream genes. Cassettes in the opposing direction cause transcriptional blocking of downstream genes. The pJQ200sk plasmid possesses the multiple cloning site from the pBluescript II SK vector and a gentamicin resistance gene as a selectable marker. Second recombination events occur naturally at low frequency whereby the vector is excised from the genome leaving in place the recombinant DNA. Selection of this event is facilitated by the presence of the *sacB* gene on pJQ200sk which causes a suicide effect in Gram-negative bacteria when grown in the presence of 5% sucrose.

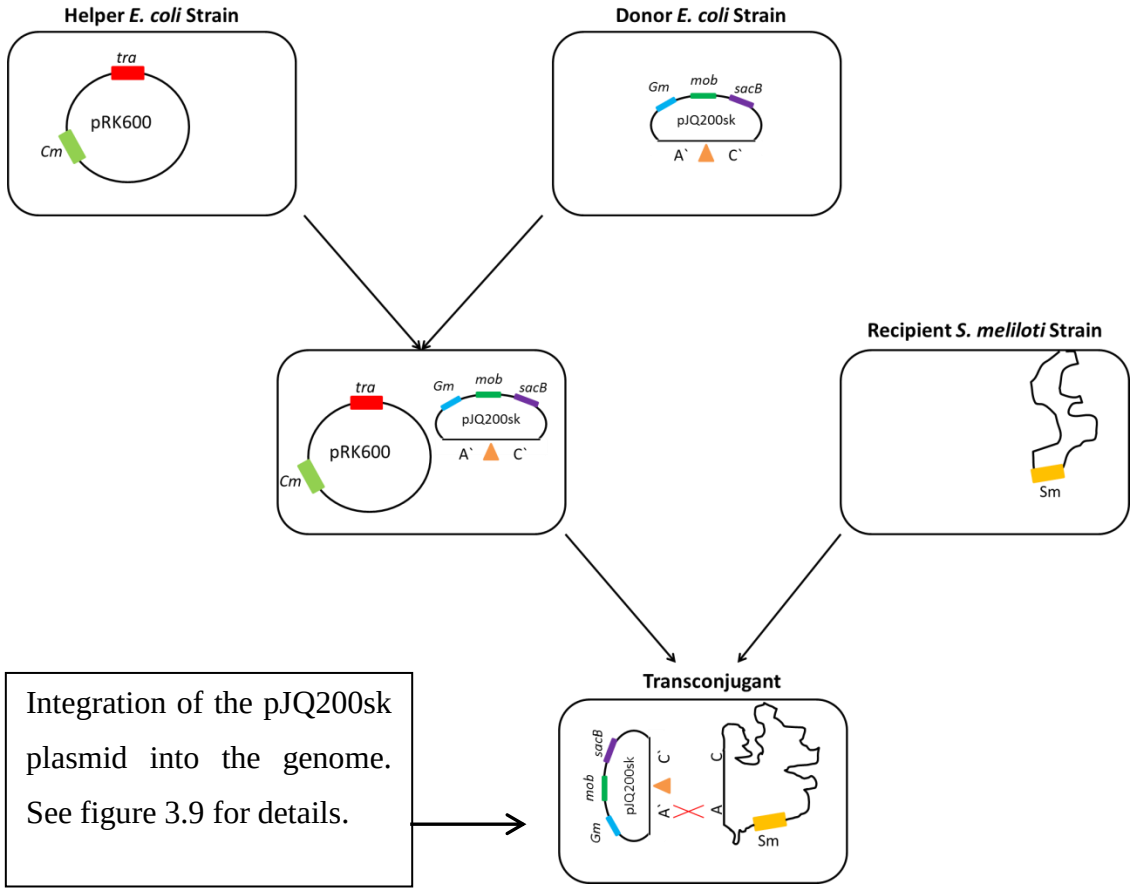


Figure 3.8: Process of triparental mating

Transfer genes (tra) are highlighted in red, mobilisation genes (mob) in green, the *sacB* gene in purple and antibiotic resistance cassettes in lime for Chloramphenicol (Cm), light blue for Gentamicin (Gm) and orange for Streptomycin (Sm). The A′,C′ and orange triangle represent the two regions of homology and the new antibiotic resistance cassette. In the helper strain the *tra* gene provides transfer functions, which encodes the elements required for conjugation. This plasmid is then transferred to the donor strain. In the intermediate stage, the donor strain contains all the elements required to transfer the plasmid of interest to the recipient, i.e. tra and mob. To ensure efficient transfer to the recipient, the transconjugant can be selected for by culturing on media containing antibiotics. Only the transconjugants with the appropriate antibiotic resistance genes (in this case, gentamicin and streptomycin) will survive. For narrow host range plasmids such as pJQ200sk, once the plasmid has been delivered to the recipient, it can be secured and maintained by integration onto the chromosome by homologous recombination (figure 3.9).

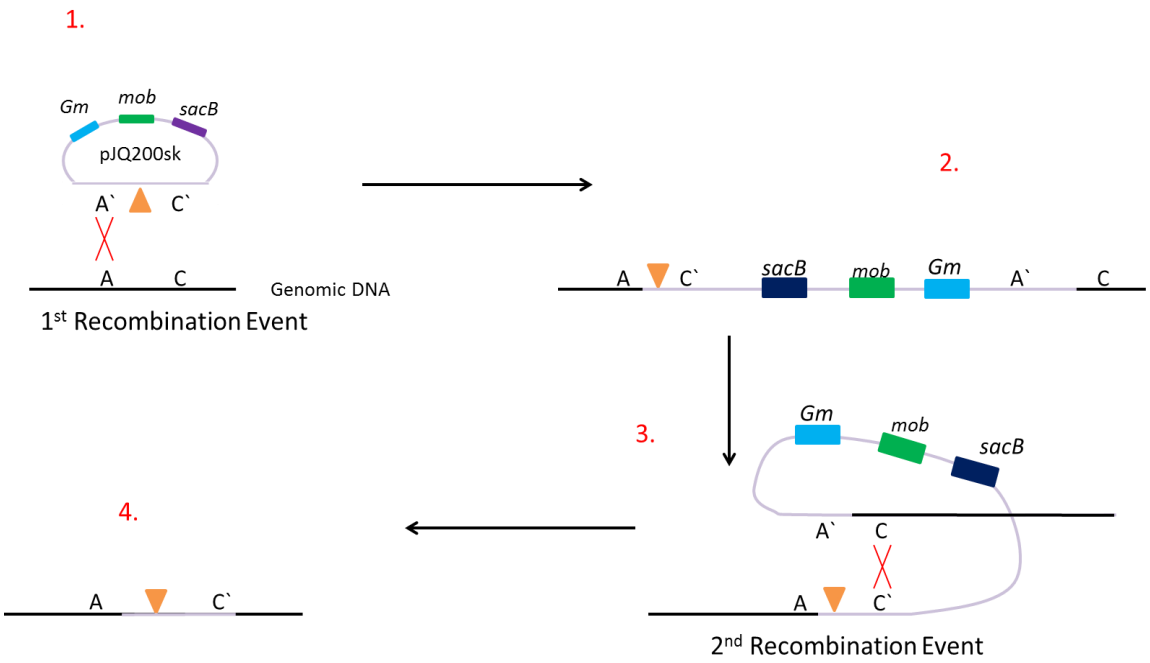


Figure 3.9: Schematic of the homologous recombination events

Circular suicide vector pJQ200sk is shown at stage 1 with antibiotic resistance gene (light blue), *sacB* gene (dark blue) and cloned genomic DNA (A′ and C′). The antibiotic resistance gene to be inserted is represented by an orange triangle. The straight black line indicates a section of chromosomal DNA with parental allele (A and C). Recombination events are marked by a red X. Stage 2 represents the entire vector inserted into the genome between the A and C positions as a result of the 1st recombination event. Stage 3 is the 2nd recombination event where the C′ allele from the vector recombines with the C allele on the genome. This leads to the excision of the vector from the genome (stage 4) resulting in the new A(resistance gene)C′ locus on the genome representing the target gene with the antibiotic resistance gene located centrally.

3.4 Antibiotic resistance cassette mutagenesis of *sma2339*

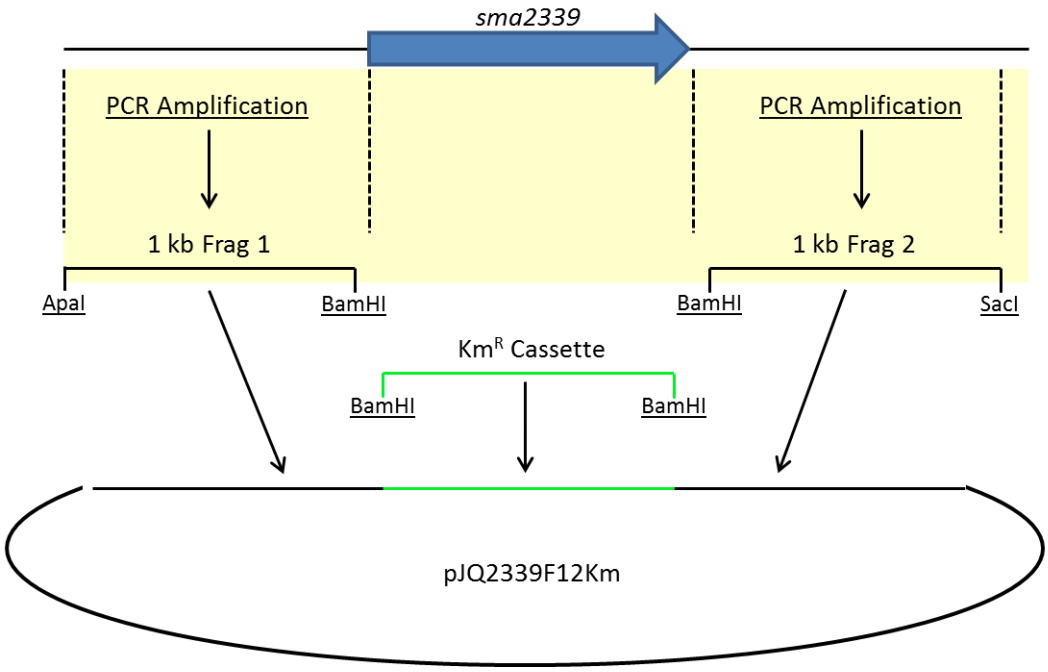


Figure 3.10: Schematic for the construction of the construct for mutagenesis of *sma2339*

A gene knockout of *sma2339* was made by replacement with a kanamycin resistance cassette. A brief overview of the process used to create the pJQ200sk based construct to mutagenize *sma2339* is shown in figure 3.10. The region upstream and downstream of *sma2339* was assessed for restriction sites amenable for the introduction of the kanamycin cassette. Two 1 kb regions were designed that would result in the complete removal of sma2339 from the genome to ensure that no residual gene activity could occur and the primer sequences for both primer pairs are given in table 3.14. A BamHI restriction site was introduced at the 3' end of the upstream fragment and on the 5' end of the downstream fragment so the kanamycin cassette could be inserted. The first primer pair, named sma2339 Del1F and sma2339 Del1R was designed to amplify the region 1316985-1317985 of pSymA and introduce an ApaI site at the 5' end and a BamHI site at the 3' end of the PCR product. A PCR was carried out with these primers using *S. meliloti* 2011 genomic DNA as template using the cycling conditions in table 3.15. The resulting PCR product was then purified by gel extraction and subjected to restriction digest by the ApaI and BamHI restriction enzymes. In parallel a previously prepared pJQ200sk+ vector was also digested with the ApaI and BamHI restriction enzymes. The resulting DNA fragments from these reactions were purified and used in a ligation reaction. Following transformation and screening of transformants a plasmid

with a size shift of 1 kb was identified indicating insertion of the PCR product. This was subjected to restriction analysis to confirm the correct insert. A positive clone was identified and named pJQ2339F1.

Table 3.14: Sequence of the primers used to mutagenize *sma2339*

Primer Name	Primer Sequence
sma2339 Del1F	CACCACGGGCCCCCGGTCTTTATCCATCCATCG
sma2339 Del1R	CACCACGGATCCTTAAAAAACCTTTGTCAGCGAGAC
sma2339 Del2F	CACCACGGATCCAGCGGCCGACCGCACCAGTT
sma2339 Del2R	CACCACGAGCTCCCCGCTCGGCTGTCAGCCAA
sma2339/Km For	GGCGGCCTAGCCGACGAGTCG
sma2339/Km Rev	ACTGAGAAGCTGGGCGGCAGG

Table 3.15: PCR cycling conditions for the amplification of *sma2339* mutagenesis fragments

PCR Cycling Conditions	
Annealing Temp	60 °C
Annealing Time	15 sec
Extension Time	20 sec

The second primer pair for the mutagenesis of *sma2339* named sma2339 Del2F and sma2339 Del2R shown in table 3.14 was designed to amplify the 1 kb region 1319097-1320097 of pSymA and to introduce a BamHI site on at the 5' end and a SacI site on the 3' end of the PCR product. A PCR was carried out using these primers and genomic DNA prepared from *S. meliloti* 2011 using the conditions outlined in table 3.15. The resulting PCR product was purified by gel extraction and subjected to restriction by BamHI and SacI. In parallel a plasmid preparation of pJQ2339F1 was carried out and digested using the BamHI and SacI enzymes. The resulting DNA was purified and used in a ligation reaction. Following transformation and transformant screening a plasmid with the expected 1 kb size shift was identified and confirmed to have the correct insert. The vector was named pJQ2339F12 and this vector was the background for the introduction of the kanamycin cassette into the BamHI site.

The kanamycin resistance cassette was sourced from the vector pUC4K. A plasmid preparation was carried out on the pUC4K vector which was subsequently subjected to a restriction digest by BamHI. In parallel the pJQ2339F12 plasmid was also digested with the BamHI restriction enzyme. A ~1.2 kb fragment representing the kanamycin resistance cassette was purified by gel extraction from the pUC4K digest along with the

~7.4 kb fragment representing the linearized pJQ2339F12 vector. These DNA fragments were then used as template for a ligation reaction and transformed into an *E. coli* host. The cells were plated onto LB agar containing kanamycin and gentamicin and colonies were screened for the presence of the kanamycin cassette. A correct plasmid was identified and named pJQ2339F12km.

The pJQ2339F12km vector was introduced into *S. meliloti* 2011 by triparental mating and first recombinants selected on TY agar containing 1000 μg/ml streptomycin and 80 μg/ml gentamicin. A number of first recombinants were picked and inoculated into 5 ml of TY broth each and allowed to grow to early stationary phase. To select for second recombination events a volume from these cultures was plated onto TY agar containing 5% sucrose and 100 μg/ml kanamycin. Colonies were individually tested for kanamycin resistance and gentamicin sensitivity. A potential mutant was identified, named S. meliloti 2011sma2339 and subjected to a final confirmation by PCR.

Genomic DNA was prepared from a culture of the newly created *S. meliloti* 2011*sma2339* and *S. meliloti* 2011. Using the primer set sma2339/Km For/Rev two PCRs were carried out using *S. meliloti* 2011 and the new mutant genomic DNA as templates. The PCR products were analysed by agarose gel electrophoresis on a 1.2% gel. The primer set amplifies the region 1317905-1319197 of pSymA resulting in a 1296 bp fragment for wild type and a 1447 bp fragment for the mutant. Figure 3.11 is the agarose gel image comparing the PCR product from the wild type to the PCR product from the new mutant. A size shift was observed in the mutant representing the kanamycin cassette at the correct position.

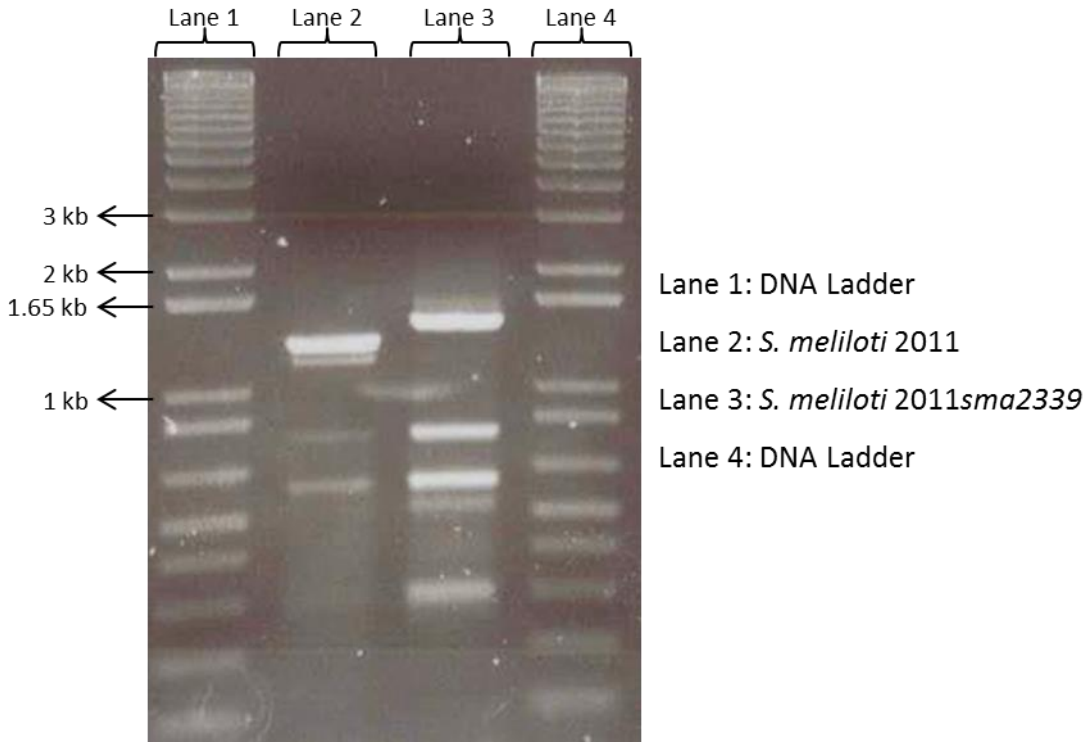


Figure 3.11: Confirmation of the *S. meliloti* 2011*sma2339* mutant by PCR

Based on the size of the top bright band in lanes 2 and 3 the kanamycin cassette is present in the region targeted by the PCR. As the secondary bands could not be removed by optimising the PCR a secondary confirmation was carried out. The top bright band in lanes 2 and 3 was gel excised and subjected to analysis by restriction digest. A restriction digest by SmaI results in different banding patterns for the wild type PCR and the mutant PCR products. The expected sizes are 118 bp, 299 bp and 876 bp for wild type and two possibilities for the mutant that are 924 bp and 520 bp or 904 bp and 540 bp depending on the cassette orientation. The result of this analysis is shown in figure 3.12. The digest for the wild type in lane 3 is an incomplete digest however it clearly shows the 876 bp and 299 bp bands. The 118 bp band is marked as it is difficult to see due to the gel image having been manipulated for this document. Lane 5 shows the two bands for the mutant genotype. It was not possible to determine the orientation of the cassette as the fragment sizes for negative and positive were similar.

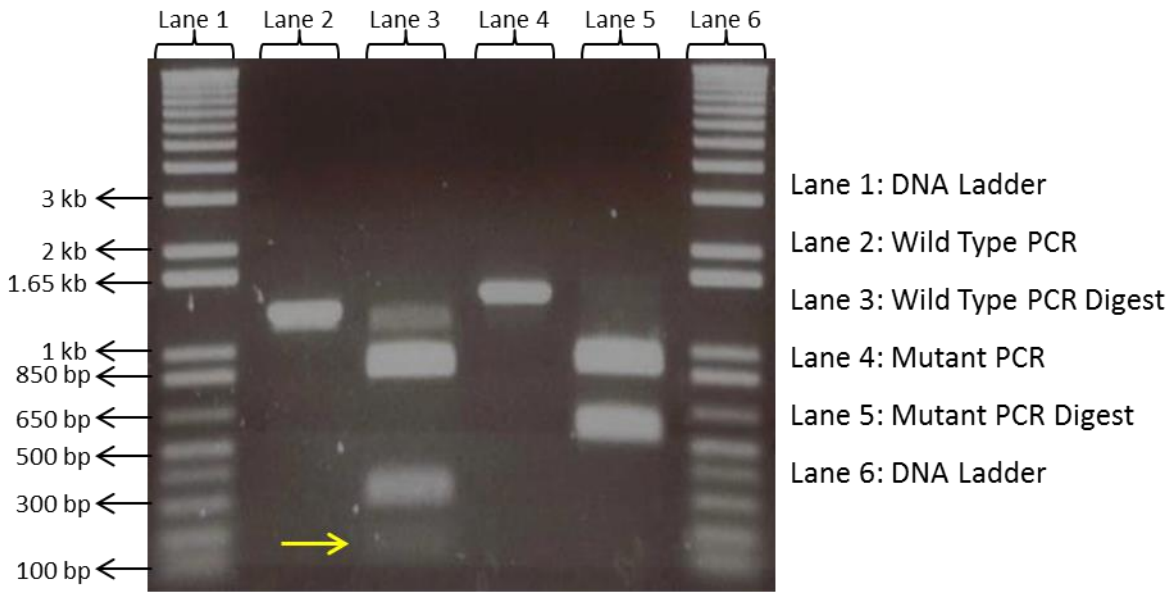


Figure 3.12: Restriction digest confirmation of the *S. meliloti* 2011*sma2339* mutant

This mutant was assessed for its ability to produce a siderophore by CAS plate assay. It was also tested for its ability to utilise rhizobactin 1021 as an iron source. It was assessed for rhizobactin 1021 utilisation to ensure no changes were introduced into *rhtA* during the mutagenesis process. This mutant tested positive for siderophore production and utilisation of rhizobactin 1021 and these results, along with the analysis methods, will be discussed in further detail in section 3.5.

After confirming the genotype of the new *S. meliloti* 2011*sma2339* mutant a complementing clone was made using the pBBR1MCS-5 plasmid containing the genes *rhtA-sma2339*. Both of these genes were cloned into the vector as they appear to be transcribed as a single transcript. To ensure that *sma2339* was transcribed it was decided to keep the genomic arrangement of the genes.

3.4.1 Cloning of *rhtA-sma2339* for the complementation of the *S. meliloti* 2011*sma2339* mutant

The *rhtA-sma2339* genes including the ribosome binding site of *rhtA* were cloned into the broad host range vector pBBR1MCS-5. Two primers were designed to amplify the region from genomic DNA prepared from *S. meliloti* 2011. The forward primer was designed to include the ribosome binding site upstream of the start codon of *rhtA* and to incorporate a KpnI at the 5' end of the PCR product. The reverse primer was designed to include the stop codon of *sma2339* and to incorporate a SacI site at the 3' end of the PCR product. The sequences of the each primer are given in table 3.16. A PCR was carried out using the conditions detailed in table 3.17 to amplify the described region from *S. meliloti* 2011 genomic DNA. The PCR product was purified by gel extraction and subsequently digested with the KpnI and SacI restriction enzymes. Concurrently the pBBR1MCS-5 vector was digested with the KpnI and SacI enzymes. The resulting DNA fragments were then purified and used in a ligation reaction. Following transformation a plasmid screen was carried out for a ~3.4 kb size shift in the vector. This clone was subsequently analysed to confirm the correct insert and named pCC101.

Table 3.16: Primer sequences for cloning *rhtA-sma2339* into pBBR1MCS-5

Primer Name	Primer Sequence
rhtA/sma2339 For	CACCACGGTACCTCTGGTGCGCAGGGGGCG
rhtA/sma2339 Rev	CACCACGAGCTCTCATGAGTTGCTCCTGTCGGGATCC

Table 3.17: PCR conditions for amplification of *rhtA-sma2339*

PCR Cycling Conditions	
Annealing Temperature	62 °C
Annealing Time	15 sec
Extension Time	2 min

The plasmid pCC101 was introduced into *S. meliloti* 2011*sma2339* by triparental mating.

The phenotype of both the mutant and the complemented mutant were assessed for the production of siderophore by the chrome azurol sulphonate, CAS plate assay.

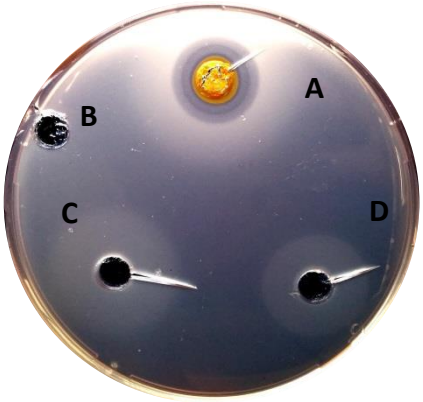
3.5 Assessment of rhizobactin 1021 utilisation by the *S. meliloti* 2011*sma2339* mutant

To ensure that the *rhtA* gene was not disrupted by the knockout of *sma2339* an iron nutrition bioassay was carried out.

To set up an iron nutrition bioassay, molten agar is seeded with the strain of interest and the iron chelator 2,2'dipyridyl is added to create an iron deplete environment. This ensures that in the absence of an exogenous iron source or siderophore, growth occurs slowly and also results in the up regulation of iron responsive genes. The plate is then poured and the agar allowed to set. Wells are aseptically pierced in the medium to allow for the test solutions to be added. The test solutions are generally a siderophore and a number of controls: $FeCl_3$ as a positive control and Fe free H_2O as a negative control with a second siderophore occasionally added as a third control. After incubation for 24-48 hours at 30 °C, haloes of intense growth appear around wells where the test solutions lead to additional growth. Growth around the positive control indicates that the inoculum is viable. Timing of the incubation is crucial when working with siderophore producing strains as background growth will occur that will occlude the haloes. Utilisation is deemed to occur if a halo appears around a well.

Previous studies have optimised the 2,2'dipyridyl concentration for *S. meliloti* 2011 to 300 μM and the production of rhizobactin was achieved by growing wild type in an iron deplete environment, filter sterilising the medium and concentrating it to one tenth of its original volume. The optimal concentration of commercially available xenosiderophores has also been previously identified (O Cuiv et al. 2008; O Cuiv et al. 2004; Lynch et al. 2001).

As it is difficult to image the siderophore bioassay, figure 3.13 is an example of a clear assay carried out with the siderophore deficient strain *S. meliloti* 2011*rhbA62* (O Cuiv et al. 2007). In the original photo of a bioassay it is difficult to observe the haloes as the contrast is not sufficient, resulting in the image having to be digitally enhanced.



A: $FeCl_3$ Control C: Ferrioxamine B

B: Fe- H_2O Control D: Coprogen

Figure 3.13: Example of a positive iron nutrition bioassay result

The test solutions in figure 3.13 are not related to the work carried out as part of this study and are only included as they demonstrate the clarity of the technique.

An iron nutrition bioassay was set up to assess the *S. meliloti* 2011*sma2339* mutant for the ability to utilise rhizobactin 1021. This was tested as the kanamycin cassette was inserted immediately downstream of *rhtA* and may have disrupted the gene in the mutagenesis process. A 200 µl aliquot of an early stationary phase culture was inoculated into 25 ml of molten agar at 45 °C. The iron chelator 2,2'dipyridyl was added to a final concentration of 300 µM. The agar was poured and allowed to set. As a test of the rhizobactin 1021 preparation, a bioassay was also set up using the *S. meliloti* 2011*rhbA62* strain. *S. meliloti* 2011*rhtA* was assessed as a negative control as it is deficient in rhizobactin 1021 utilisation. Wells were cut in the agar and the following solutions added: $FeCl_3$, Fe free H_2O, concentrated supernatant from a wild type deplete culture (rhizobactin +) and concentrated supernatant from a wild type replete culture (rhizobactin -). The results of this analysis are shown in table 3.18.

Table 3.18: Analysis of rhizobactin 1021 utilisation by *S. meliloti* 2011*sma2339*

Strain	$FeCl_3$	Fe- H_2O	rhizobactin -	rhizobactin +
S. meliloti* 2011*rhbA62	+	-	-	+
S. meliloti* 2011*rhtA	+	-	-	-
S. meliloti* 2011 *sma2339	+	-	-	+

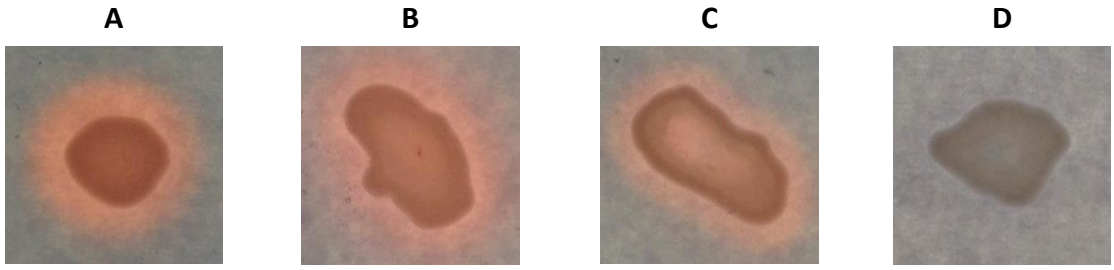
The *S. meliloti* 2011*sma2339* is positive for rhizobactin 1021 utilisation showing the mutation did not interfere with *rhtA*. The controls for the experiment showed that there

was no growth enhancing effect from the concentrated supernatant from a wild type replete culture, confirming that rhizobactin 1021 was responsible for the growth.

3.6 Assessment of siderophore production by the *S. meliloti* 2011*sma2339* mutant

As the *sma2339* gene is possibly involved in the production of rhizobactin 1021, assessing siderophore production from *S. meliloti* 2011*sma2339* will determine if it is essential or redundant to this process. The method for assessing siderophore production is the chrome azurol sulphonate agar plate assay and is carried out using the method described in section 2.10.1. In brief a single colony of a *S. meliloti* strain is inoculated into a 5 ml aliquot of TY medium and incubated at 200 rpm at 30 °C until early stationary phase. Cultures are then normalised to an OD_{600} of 1.0 by dilution in 0.85% saline and a 1 ml volume of this normalised culture is pelleted at 6,000 rpm for 2 minutes. The supernatant is discarded and the cell pellet suspended in 100 µl of Fe free H_2O. A 10 µl aliquot of this suspension is spotted onto a CAS agar plate and the plate incubated at 30 °C for 48 hrs. A positive result for siderophore production is indicated by a colour change in the medium from a dark rich blue colour to a reddish/orange colour. As this assay depends on factors such as diffusion and bacterial growth for a halo to form it can only be used as a qualitative measurement of siderophore production. Observation of replicate cultures on different batches of CAS agar have yielded differing halo sizes (results not shown), however siderophore producing strains and non-producing strains always give a distinct halo or lack of halo respectively.

On a single CAS agar plate the strains *S. meliloti* 2011, *S. meliloti* 2011*sma2339*, *S. meliloti* 2011*sma2339* (pCC101) and *S. meliloti* 2011*rhbA62* were spotted as described previously. Growth was allowed to proceed at 30 °C for 48 hrs after which the plate was photographed. The result of this is shown in figure 3.14.



A: *S. meliloti* 2011

B: *S. meliloti* 2011*sma2339*

C: *S. meliloti* 2011*sma2339* (pCC101)

D: *S. meliloti* 2011*rhbA62*

Figure 3.14: Analysis of siderophore production by *S. meliloti* 2011*sma2339*

As shown the wild type strain (A) has a clear orange halo indicating siderophore production with the negative control (D) showing no colour change on the medium. The two test strains B and D are clearly positive for siderophore production. As the *S. meliloti* 2011*sma2339* strain is producing siderophore it demonstrates that *sma2339* is not essential for rhizobactin 1021 production in *S. meliloti* 2011. There are a number of possible reasons for this strain retaining the ability to produce siderophore:

- The Sma2339 protein is involved in rhizobactin 1021 biosynthesis but can be complemented by the RhbD protein resulting in an intact rhizobactin 1021.
- The Sma2339 protein is involved in biosynthesis and is complemented by RhbD resulting in an altered form of rhizobactin 1021 possibly lacking a lipid structure.
- The Sma2339 protein is not directly involved in biosynthesis but plays a non-essential periphery role in production.

To address the reason for the *S. meliloti* 2011*sma2339* mutant retaining the ability to produce siderophore a mutation was constructed that knocks out the *rhbD* gene. This mutant was made in two backgrounds, *S. meliloti* 2011 and in *S. meliloti* 2011*sma2339*. A *S. meliloti* 2011*rhbD* mutant could be assessed for the ability of Sma2339 to complement *rhbD* because if rhizobactin 1021 is produced from this strain then the Sma2339 protein is most likely complementing the RhbD protein. In the case that siderophore is produced by *S. meliloti* 2011*rhbD* it must be confirmed that the Sma2339 protein is responsible. The *S. meliloti* 2011*sma2339rhbD* mutant was assessed by

comparison with *S. meliloti* 2011*rhbD* whether the Sma2339 protein is the complementing protein for RhbD.

In addition to assessing siderophore production it was subsequently determined whether an altered structure was produced. This was of particular interest when comparing the individual mutants *S. meliloti* 2011*sma2339* and *S. meliloti* 2011*rhbD* with the double mutant and the wild type.

3.7 Antibiotic resistance cassette mutagenesis of *rhbD*

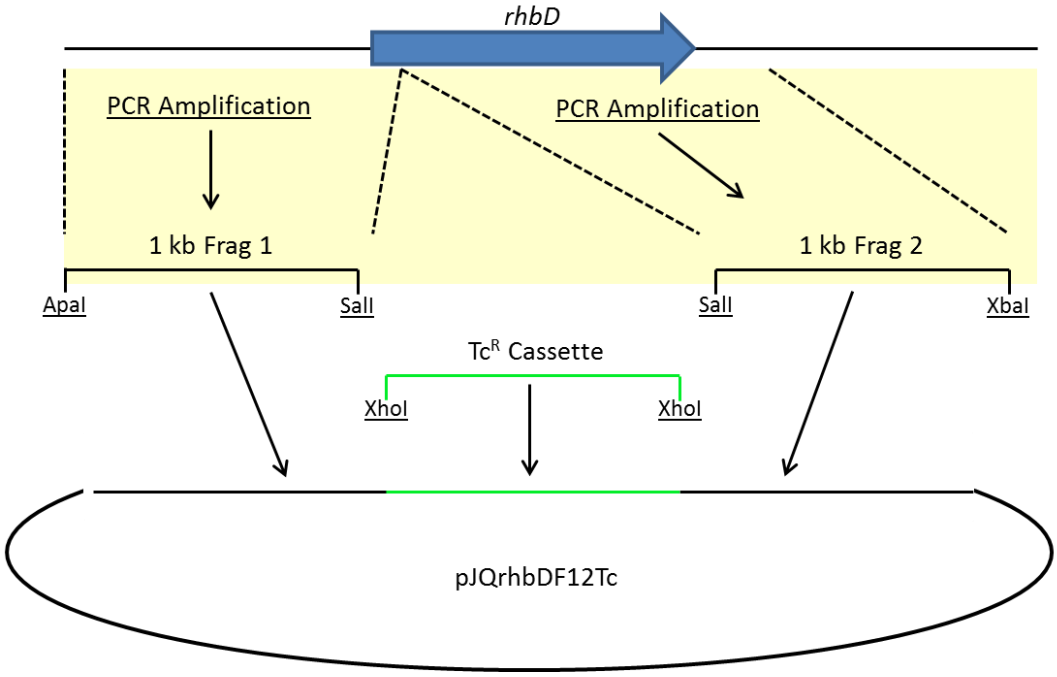


Figure 3.15: Schematic of the construction of the construct for mutagenesis of *rhbD*

A gene knockout of *rhbD* was made by disruption with a tetracycline resistance cassette. A brief overview of the process used to create the pJQ200sk based construct to mutagenize *rhbD* is shown in figure 3.15. The region internal to and surrounding *rhbD* was assessed for restriction sites amenable for the introduction of a tetracycline resistance cassette and no naturally occurring sites were present. As a result of this lack of restriction sites it was decided to introduce the tetracycline resistance cassette into an artificially introduced restriction site 8 bp downstream from the start codon at position 1310820 of the pSymA megaplasmid. Two primer pairs, shown in table 3.19 were designed to amplify the region for mutagenesis and to introduce a SalI site at position 1310820. The first primer pair named rhbDmut1F and rhbDmut1R was designed to

amplify the 1 kb region 1309821-1310820 of pSymA and to introduce an ApaI site at the 5' end and a SalI site on the 3' end of the PCR product. A PCR was carried out with these primers using *S. meliloti* 2011 genomic DNA as template with the cycling conditions in table 3.20. The resulting PCR product was then purified by gel extraction and subjected to restriction digest by the ApaI and SalI restriction enzymes. The pJQ200sk+ vector was also digested with ApaI and SalI restriction enzymes. The resulting DNA from these reactions was purified and used in a ligation reaction. Following transformation and screening of transformants a plasmid with a size shift of 1 kb was identified indicating insertion of the PCR product. This was subjected to restriction analysis to confirm the insert was correct and in the right orientation. A positive clone was identified and named pJQrhbDF1.

Table 3.19: Sequence of the primers used to mutagenize *rhbD*

Primer Name	Primer Sequence
rhbDmut1F	CACCACGGGCCCGGCGAGGCCGGCAAGCCATGG
rhbDmut1R	GAGGAGGTCGACCTCAAGCATGTGCCAGAACCTTTTCCG
rhbDmut2F	CACCACGTCGACGAGGAGCGAGCCGGACACTGCC
rhbDmut2R	CACCACTCTAGAACCACTTCCTCGCCGAAGCGGCACG
Tetcas For	CTAGTAACGGCCTCGAGTGTGCTGG
Tetcas Rev	CCGCCAGTGTGCTCGAGATCTGCAGAA
rhbDTc For	ATGCTGGCGCCGACGCTGC
rhbDTc Rev	CATCTTCCATTGCGGCACG

Table 3.20: PCR cycling conditions for amplifing *rhbD* mutagenesis fragments

PCR Cycling Conditions	
Annealing Temp	62 °C
Annealing Time	15 sec
Extension Time	30 sec

The second primer pair for the mutagenesis of *rhbD* named rhbDmut2F and rhbDmut2R shown in table 3.19 was designed to amplify the ~1 kb region 1310820-1311760 of pSymA and to introduce a SalI site at the 5' end and a XbaI site at the 3' end of the PCR product. A PCR was carried out using these primers on genomic DNA prepared from *S. meliloti* 2011 using the conditions outlined in table 3.20. The resulting PCR product was purified by gel extraction and subjected to restriction by the SalI and XbaI restriction enzymes. In parallel a freshly prepared pJQrhbDF1 plasmid was digested with SalI and

XbaI. The resulting DNA fragments were purified and used as the template in a ligation reaction. Following transformation and transformant screening a plasmid with the expected size shift of ~1 kb was identified. This was subjected to restriction analysis to confirm that the insert was correct and a positive clone was identified. This vector was named pJQrhbDF12 and this is the vector into which the tetracycline cassette was inserted using the SalI site.

The tetracycline resistance cassette was sourced from the vector pHP45ΩTc. Primers were designed to amplify the tetracycline cassette from the vector and add an XhoI restriction site to each end of the PCR product. These primers were named Tetcas For and Tetcas Rev and the sequences are shown in table 3.19. A PCR was carried out using these primers and the pHP45ΩTc vector as template and a prominent ~2 kb band representing the tetracycline cassette was purified by gel extraction. This fragment was digested with the XhoI restriction enzyme with the pJQrhbDF12 being digested with SalI in parallel. The XhoI and SalI restriction enzymes create compatible ends which allowed for the insertion of the tetracycline cassette into the vector. The digested restrictions were purified and used as the template for a ligation reaction. Following transformation and plating onto LB agar containing tetracycline and gentamicin a plasmid screen was carried out. A correct plasmid was identified and named pJQrhbDF12Tc.

The pJQrhbDF12Tc vector was introduced in *S. meliloti* 2011 and *S. meliloti* 2011*sma2339* by triparental mating and first recombinants selected for on agar containing 1000 μg/ml streptomycin and 80 μg/ml gentamicin. A number of first recombinants for each strain were picked and inoculated into 5 ml of TY broth and growth allowed to proceed to early stationary phase. To select for second recombination events a volume from these cultures was plated onto TY agar containing 5% sucrose and 2.5 μg/ml tetracycline. Colonies were individually screened for tetracycline resistance and gentamicin sensitivity. A potential mutant was identified in each of the background strains *S. meliloti* 2011 and *S. meliloti* 2011*sma2339*, named *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* respectively, and subjected to a final confirmation by PCR.

Genomic DNA was prepared from a culture of both of these newly created mutants. Using the primers rhbDTc For and rhbDTc Rev four PCRs were carried out using

genomic DNA extracted from *S. meliloti* 2011, *S. meliloti* 2011I*rhbD* and *S. meliloti* 2011*sma2339rhbD* as templates. The PCR products were analysed by agarose gel electrophoresis on a 0.7% agarose gel. The rhbDTc For and rhbDTc Rev primer pair amplifies the region 1310670-1310970 which contains the insertion site of the tetracycline cassette. The wild type PCR results in a 300 bp PCR product with a positive result for the mutants resulting in a 2.2-2.3 kb fragment. Figure 3.16 shows the bands obtained from the agarose gel electrophoresis confirmation.

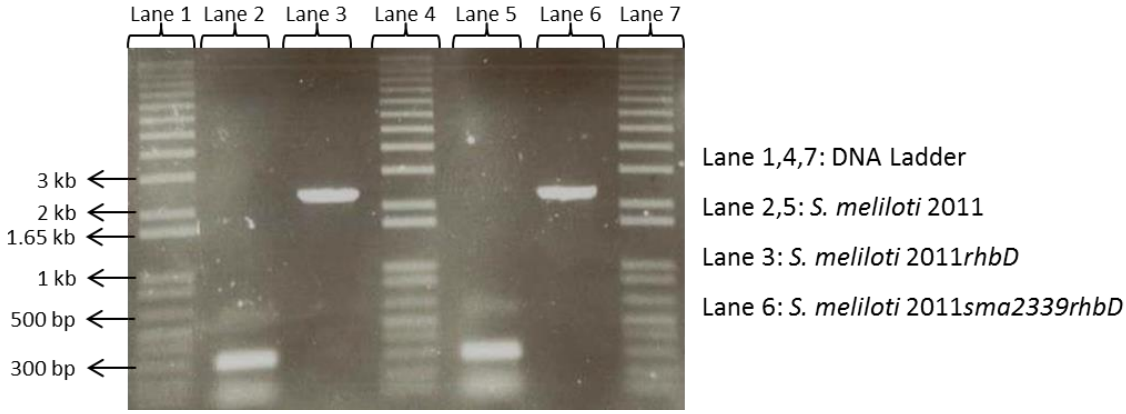


Figure 3.16: Confirmation of the *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* mutants by PCR

As can be seen in figure 3.16 a clear size shift was observed in both mutants indicating insertion of the tetracycline cassette at the correct position. Lane 3 and lane 6 are the PCR products attained from *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* respectively.

In parallel to creating these mutants two complementing clones were constructed. It was necessary to make two complementing clones as the method used to knockout *rhbD* will result in the knockout of *rhbE* and *rhbF* due to the polar effect. The first complementing clone introduced three genes *rhbDEF* while the second introduced just *rhbEF*. The *rhbDEF* complementation should recover the background phenotype in both mutants and show that it is possible to restore siderophore production with three of the biosynthesis genes on a plasmid. The second complementing clone will allow the effect of the abolished *rhbD* gene to be determined.

3.7.1 Cloning of *rhbDEF* and *rhbEF* for the complementation of the *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* mutants

The *rhbDEF* genes were cloned into the broad host range vector pBBR1MCS-5. Two primers were designed to amplify the genes from the genome of *S .meliloti* 2011. The forward primer was designed to include the ribosome binding site of *rhbD* and introduce an ApaI site at the 5' end of the PCR product. The reverse primer was designed to include the stop codon of *rhbF* and introduce an XbaI site onto the 3' end of the PCR product. The primers were named rhbDEF For and rhbF Rev and the sequences are given in table 3.21. Using the conditions outlined in table 3.22 the *rhbDEF* genes were amplified from *S. meliloti* 2011 genomic DNA. The PCR product, ~3.7 kb, was purified by gel extraction and the purified DNA fragment subjected to a restriction digest by ApaI and XbaI. A preparation of pBBR1MCS-5 plasmid was also digested with ApaI and XbaI. Both DNA fragments were purified and used in a ligation reaction. Following transformation a plasmid screen was carried out and a clone of correct size was identified. This clone was confirmed to be correct and was named pCC102. This clone was then introduced into *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* to complement the *rhbD* mutant phenotype.

Table 3.21: Primer sequences for the amplification of *rhbDEF* and *rhbEF* for cloning into pBBR1MCS-5

Primer Name	Primer Sequence
rhbDEF For	CACCACGGGCCCGGAAAAGGTTCTGGCACATGCTTGAG
rhbEF For	CACCACGGGCCCTTCATTCAACAATTCGGATTATGATCATGA
rhbF Rev	CACCACTCTAGATTACGCGGCGTTGCTGCGCTTGTCG

Table 3.22: PCR cycling conditions for the amplification of *rhbDEF* and *rhbEF*

PCR Cycling Conditions	
Annealing Temp	62 °C
Annealing Time	15 sec
Extension Time	105 sec

The *rhbEF* genes were also cloned into the pBBR1MCS-5 vector. A primer was designed to include the ribosome binding site of *rhbE* and incorporate an ApaI site at the 5' end of the PCR product. This was named rhbEF For and the sequence is given in table 3.21. The reverse primer used was rhbEF Rev and is described previously. Using

the conditions described in table 3.22 a PCR was carried out using the rhbEF For and rhbF Rev primers and *S. meliloti* 2011 genomic DNA as template. The resulting ~3.1 kb fragment was purified by gel extraction and subjected to restriction digest with ApaI and XbaI along with the pBBR1MCS-5 vector. The resulting fragments from these digests were purified and used in a ligation reaction. This reaction was transformed into *E. coli* and a plasmid screen carried out to identify positive clones. A clone was confirmed as positive and named pCC103. This clone was introduced into *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* by triparental mating to allow for the *rhbD* deficient phenotype to be assessed.

3.8 Assessment of siderophore production by the *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD* mutants

The following strains were assessed for their ability to produce siderophore: *S. meliloti* 2011, *S. meliloti* 2011*rhbD*, *S. meliloti* 2011*rhbD* (pCC102), *S. meliloti* 2011*rhbD* (pCC103), *S. meliloti* 2011*sma2339rhbD*, *S. meliloti* 2011*sma2339rhbD* (pCC102), *S. meliloti* 2011*sma2339rhbD* (pCC103) and *S. meliloti* 2011*rhbA62*. The *S. meliloti* 2011 and *S. meliloti* 2011*rhbA62* strains were included as positive and negative controls respectively. Each strain was grown to early stationary phase in TY broth containing antibiotics where appropriate, normalised as described previously and suspended in 100 ul Fe- H_2O. To ensure strong expression from the plasmid encoded genes 100 µl of 100 mM IPTG was spread onto each plate to induce their expression. A 10 µl volume of each strain was spotted onto CAS agar plates containing IPTG and growth allowed to proceed for 48 hrs at 30 °C until a halo was clearly visible around the positive control. The results of this experiment are shown in figure 3.17.

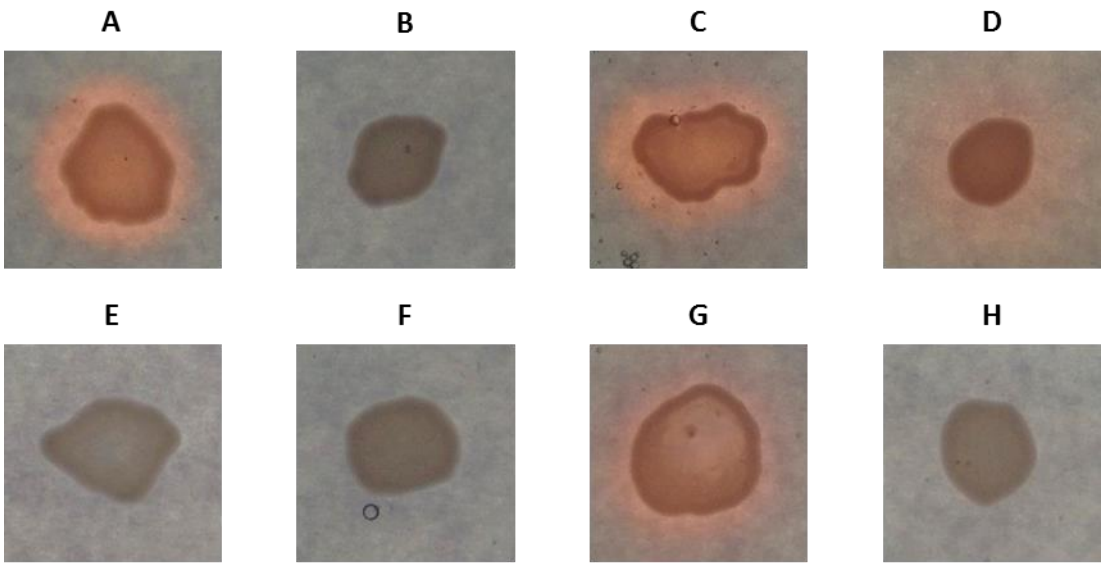


Figure 3.17: Assessment of siderophore production from *S. meliloti* 2011*rhbD* and *S. meliloti* 2011*sma2339rhbD*

The positive control A shows a clear orange colouration of the medium indicating siderophore production with the negative control E showing no orange colouration of the medium. As expected with the tetracycline cassette insertion in *rhbD* there is no siderophore production as shown in cultures B and F. This is expected as the cassette insertion will result in the polar knockout of *rhbEF* which would definitely result in abolished siderophore production. The mutants were complemented with pCC102, encoding *rhbDEF*, to show that the biosynthesis operon can be complemented and cultures C and G clearly show siderophore production from the complemented strains. The mutants were complemented with pCC103, encoding *rhbEF*, to assess the *rhbD* negative phenotype. Culture D demonstrates that siderophore production is still possible without the RhbD protein in a wild type background. Culture H represents a strain that is deficient in the acetyltransferases *rhbD* and *sma2339* and the negative result indicates that at least one acetyltransferase is required to produce a functional siderophore.

The results presented here, along with the results from section 3.6 demonstrate that siderophore production is still possible when only one of the acetyltransferases is present. However when both acetyltransferases are disrupted it results in no siderophore

production. This shows that both RhbD and Sma2339 can complement each other's activity however Sma2339 does not appear to be fully able to complement RhbD. This is apparent as there is no halo in culture D in figure 3.17 rather it is just a discolouration of the area occupied by the culture. This suggests that RhbD is the primary acetyltransferase for the biosynthesis of rhizobactin 1021 due to the minimal siderophore production in the RhbD deficient strain.

These siderophore production experiments have allowed for essential genes to be identified however they yield no information on the structure of the produced siderophore. To ascertain the structure of the produced siderophore from each strain extracts were then taken from each producing strain and analysed by electrospray ionisation mass spectrometry (ESI-MS).

3.9 Analysis of mutant produced siderophores by ESI-MS

3.9.1 Sample preparation and siderophore concentration determination

The siderophore producing strains *S. meliloti* 2011, *S. meliloti* 2011*sma2339*, *S. meliloti* 2011*rhbD* (pCC103) and the non-producing strain *S. meliloti* 2011*rhbA62* were selected for analysis by ESI-MS. A single colony of each strain was inoculated into 5 ml of TY broth, incubated at 200 rpm at 30 °C and growth allowed to proceed to early stationary phase. A 1 ml aliquot of this culture was inoculated into 100 ml of Trizma medium supplemented with 200 µM 2'2,dipyridyl. For the *S. meliloti* 2011*rhbD* (pCC103) strain IPTG was included at 200 µM along with gentamicin at a final concentration of 60 µg/ml. Cultures were grown for 48 hr to maximise the concentration of siderophore produced. The cells were centrifuged at 6,000 rpm for 15 min to pellet the cells and the supernatant collected. The siderophores were extracted by the method described in section 2.11.2 using XAD-4 resin. After extraction and suspension of the extracts in 6 ml of molecular grade H_2O, the level of extracted siderophore was measured. It was decided to measure the siderophore levels after extraction as it was found that the spent Trizma medium interfered with the accurate detection of siderophore levels. Only a qualitative test for siderophore production was carried out prior to extraction to ensure siderophore was present.

The level of siderophore in solution was measured by the liquid CAS assay described in section 2.10.2. To determine the concentration of siderophore produced by each strain first a standard curve must be created. This was achieved by using the commercially available siderophore ferrioxamine B that allowed standard solutions to be made. These solutions were then tested by the liquid CAS assay and a standard curve drawn. A 500 µl volume of liquid CAS was placed in a microfuge tube to which 100 µl of the siderophore solution was added and the volume made up to 1 ml. To this solution 10 µl of a 200 µM 5-sulfosalicyclic acid solution, known as shuttle solution was added and the mix incubated at 37 °C for 30 min to allow the reaction run to completion. The samples absorbance was read at OD_{630} which detects the level of unreacted blue CAS dye, as a result the assay must be blanked with CAS dye that is overloaded with siderophore and this was achieved by using 100 µl of a 100 mM solution of desferrioxamine B. Figure 3.18 is the standard curve attained with the CAS assay, it

does not form a linear relationship to the siderophore concentration however it is consistent between assays.

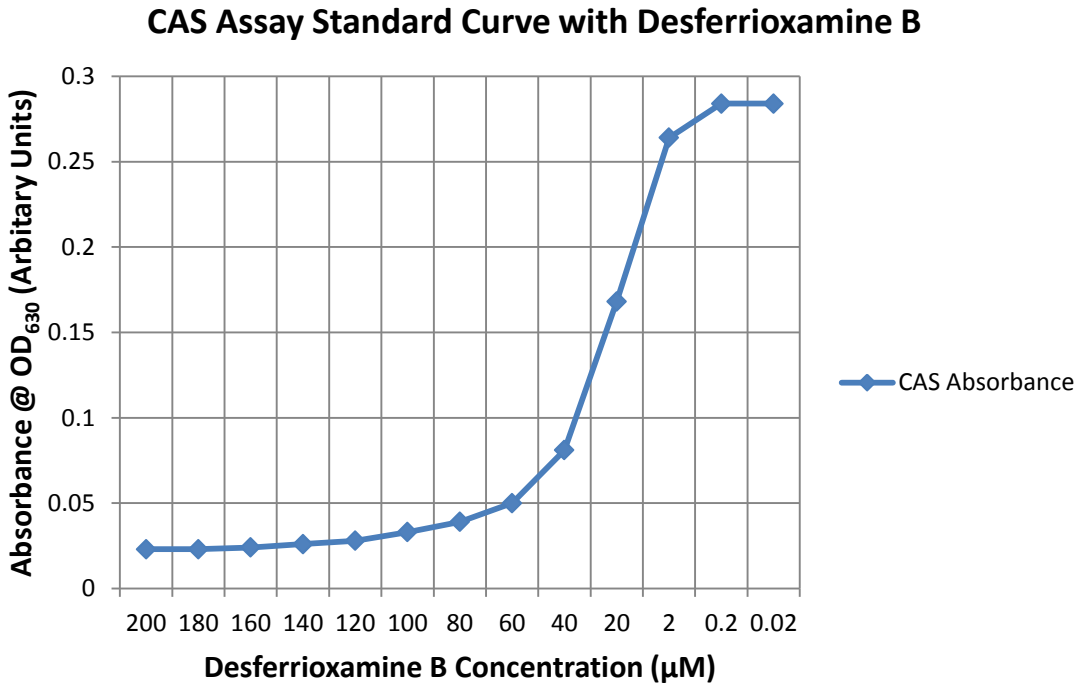


Figure 3.18: CAS Assay Standard Curve with Desferrioxamine B

Using the standard curve in figure 3.18 the concentration of the extracted siderophore for each strain was determined. Various dilutions of the extracted siderophore samples were carried out to obtain OD_{630} readings that allowed for accurate readings to be attained. Using these readings the molarity and weight of the siderophore extract was calculated based on the molecular weight of rhizobactin 1021.

Table 3.23: Nominal levels of siderophore extracted from rhizobactin mutant strains

Sample Extract	Molarity	Total Yield	% of Wild Type
S. meliloti* 2011*rhbA62	0	0 g	—
***S. meliloti* 2011*rhbD* (pCC103)**	4 µM	12.72 µg	0.48 %
S. meliloti* 2011*sma2339	100 µM	318 µg	12.04 %
***S. meliloti* 2011**	833 µM	2.64 mg	100 %

As expected due to indications given by earlier results there is no siderophore produced by *S. meliloti* 2011*rhbA62* and very little siderophore produced by *S. meliloti* 2011*rhbD* (pCC103). There is a markedly reduced level of production in *S. meliloti* 2011*sma2339*

compared to wild type. The reduction in production is not entirely unexpected as these mutations have more than likely interfered with siderophore production.

3.9.2 High Pressure Liquid Chromatography of siderophore extracts

To assess the possibility of performing preparative HPLC on the siderophore extracts the siderophore samples were subjected to HPLC analysis. Preparative HPLC will allow for the exact structure of the produced siderophores to be determined by advanced instrumentation and any changes are observed in the ESI-MS analysis.

Prior to injection onto the HPLC system the siderophore extract was generally diluted either 1 in 2 or 1 in 3 to give a clear chromatogram for the samples. An Eclipse XDB-C18 HPLC column with a 5 µm particle size and dimensions 4.6 mm x 150 mm was used to achieve separation of the siderophore from contaminants in the sample. The mobile phase was methanol:H_2O at a ratio of 80:20 with a flow rate of 1 ml/minute. System pressure was maintained at ~100 Bar, with any drastic variations indicating issues with the HPLC apparatus. The typical chromatograms achieved from each sample are given in figures 3.19 to figure 3.21. The chromatogram from the *S. meliloti* 2011*rhbD* (pCC103) extract did not show a distinct peak containing siderophore as the concentration was only a small percentage of that seen in wild type.

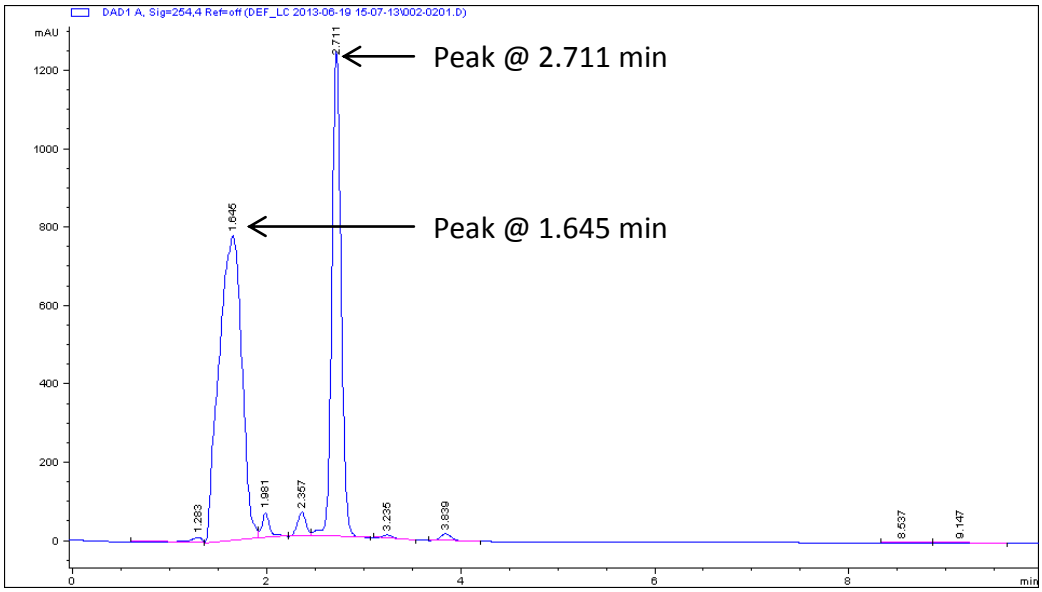


Figure 3.19: HPLC Chromatogram of siderophore extracted from *S. meliloti* 2011

The chromatogram shown in figure 3.19 obtained from the *S. meliloti* 2011 extract displays two distinct peaks the first at 1.645 min and the second more prominent peak at

2.711 min. Fraction collection of each of these peaks and subsequent analysis by the CAS assay showed the second peak to contain the entire quantity of the injected siderophore.

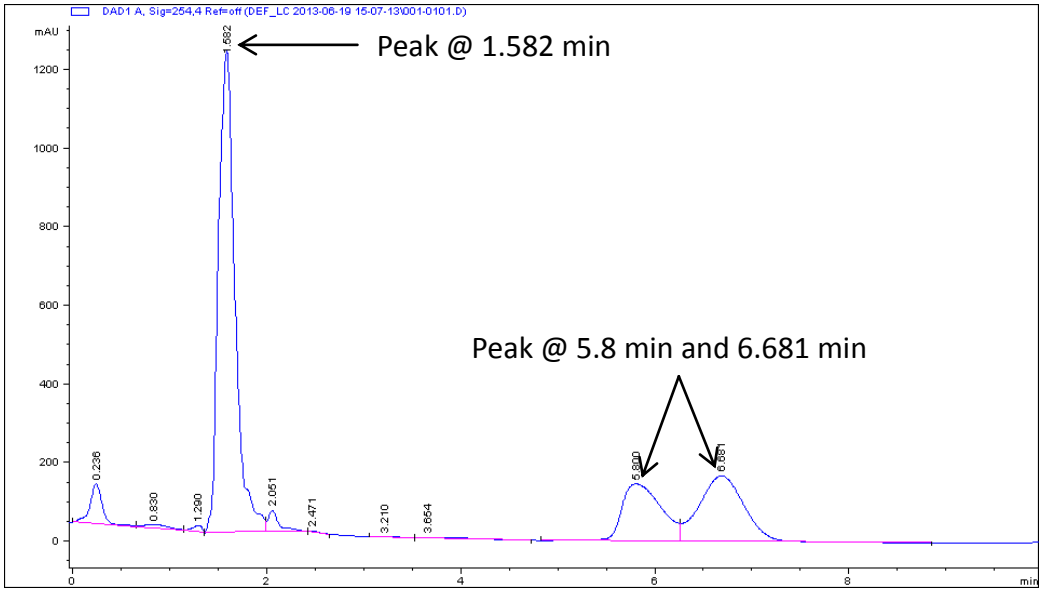


Figure 3.20: HPLC Chromatogram of siderophore extracted from *S. meliloti* 2011*rhbA62*

The chromatogram shown in figure 3.20 obtained from the *S. meliloti* 2011*rhbA62* extract displays three distinct peaks, the first at 1.582 min and the other two partially overlapping at 5.8 min and 6.681 min. Fraction collection of each of these peaks and subsequent CAS analysis detected no siderophore activity. The first peak overlays the contaminant peak found in *S. meliloti* 2011 indicating that it is probably comprises of co-eluted media constituents of the Trizma medium. The peaks at 5.8 min and 6.681 min may also represent loosely bound media constituents that cannot bind in the presence of the siderophore. These may have come through in the extraction as there would have been excess capacity on the resin due to there being no siderophore present, allowing for weakly bound molecules to come through the wash procedures.

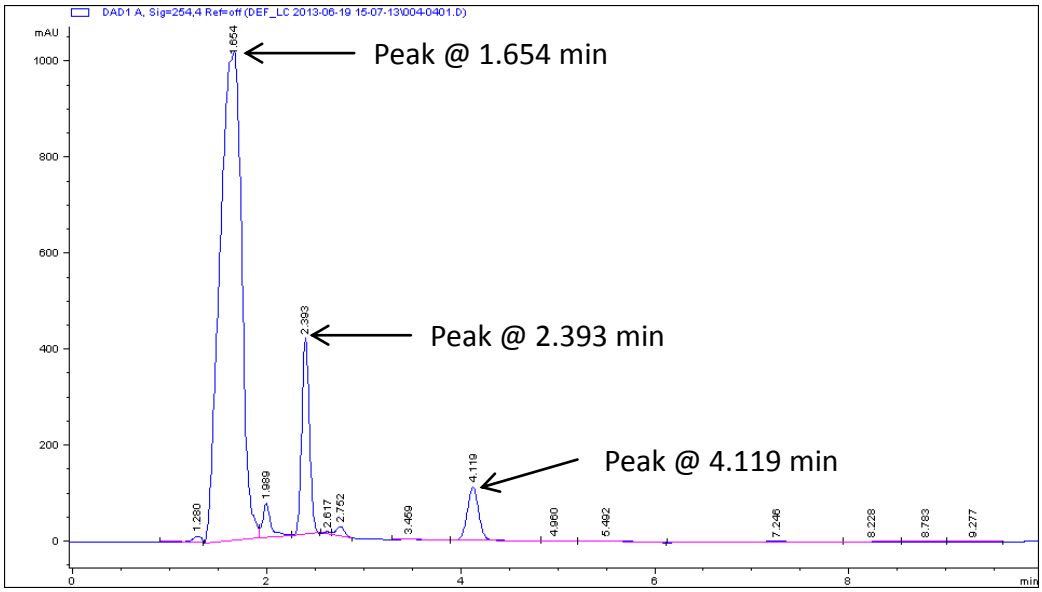


Figure 3.21: HPLC Chromatogram of siderophore extracted from *S. meliloti* 2011*sma2339*

The chromatogram shown in figure 3.21 obtained from the *S. meliloti* 2011*sma2339* extract displays three distinct prominent peaks. The first peak at time 1.654 min overlaps very closely with the main contaminant peak observed in both the *S. meliloti* 2011 and *S. meliloti* 2011*rhbA62* samples. The second peak at time 2.394 min elutes very close to the retention time of the siderophore peak in the wild type sample and CAS analysis showed this to be a siderophore containing peak. The third peak at 4.119 min was collected and tested for the presence of siderophore and no CAS reactivity was observed.

As demonstrated by these chromatograms it is possible to achieve a clean separation of the siderophore containing peak through the use of HPLC. The retention time for rhizobactin 1021 is related to the concentration of methanol in the mobile phase, in that the higher the methanol content the lower retention time. If required the methanol concentration can be lowered to allow for any possible contaminants to pass through the column therefore resulting in a high degree of purity of the siderophore. Observation of the retention time of varying methanol concentrations show that retention of the siderophore is greatly increased at a methanol concentration less than 50%.

The slight differences in the retention times for all peaks can be explained by the mobile phase and sample constituting non-buffered solutions. This can result in slight variations in retention times and changes of up to 30 sec in the retention time. Changes

were observed between replicate injections of the same siderophore preparation indicating that the slight variations in the siderophore peaks presented in figure 3.19 and figure 3.21 are within the observed tolerances. This phenomenon could not be corrected as buffering the solution with a standard buffer such as PBS will introduce metal ions and salt into the eluted siderophore preparation. Contaminating salts, in particular sodium containing salts can interfere with downstream applications such as Nuclear Magnetic Resonance (NMR) and ESI-MS to the extent that results from these techniques can become non-informative. However considering this, the inconsistencies in the retention times were not extreme enough to cause major issues with the preparative HPLC.

This method was developed as a contingency to direct ESI-MS analysis of the sample extract. The samples were first assessed by ESI-MS to ascertain whether the sample contained the intact protonated molecular mass of rhizobactin 1021 at 531 Da. If a mass change was observed by ESI-MS then preparative HPLC would be used to allow an exact structure to be determined by NMR.

3.9.3 ESI-MS analysis of siderophore extracts

ESI-MS analysis was carried out on the HCT esquire Mass Spectrometer from Bruker fitted with an electrospray ioniser. ESI-MS was ideal for this analysis in that it is a soft ionisation technique. In brief, this allows for intact molecules to be analysed providing molecular weights representative of the size and structure. It also allows ion fragmentation to be carried out in which ions are fragmented with high energy gas to give a distinct fragmentation pattern. A hard ionisation technique will not allow for intact structures to be analysed as the sample is burned at a high temperature therefore disintegrating the molecule. An example of this type of technique is Inductively Coupled Plasma Mass Spectrometry, ICP-MS, and would be useful for identifying the elemental composition of a pure sample which is not the case here.

It was important to analyse the entire sample as purifying out just the siderophore containing fraction may lead to other features being overlooked and this was achieved by analysing the sample by direct injection. However analysis of a complex sample can lead to issues when discerning peaks of interest from background noise. To address this possibility it was decided to use the extract from the *S. meliloti* 2011 sample as a

baseline against which to judge the other siderophore samples. This was done by obtaining MS^2 and MS^3 data on the fragmentation pattern of the 531 m/z peak in the *S. meliloti* 2011 sample. As it is highly unlikely that an unrelated molecule would yield an identical or highly similar fragmentation pattern in this sample set, the presence of a 531 ion with the same or highly similar fragmentation pattern will be taken as a positive indicator of rhizobactin 1021 being present in the sample.

Samples were prepared as described in section 3.9.1 and assessed for the presence of siderophore. Prior to injection formic acid was added to a final concentration of 0.1% to protonate the ions in the sample. This is required as only charged ions are detected by ESI-MS, protonation achieves charge through the addition of Hydrogen to amenable groups resulting in a mass increase of 1 Dalton and a positive charge being given to the molecule. Also, as the sample will obtain a positive charge, the ESI-MS can be carried out using the positive mode which allows for greater sensitivity of detection. The protonation procedure is also the reason that samples must be suspended in a non-buffered solution as salts and metal ions will protonate preferentially over larger molecules resulting in an extremely high background signal.

The instrument settings used for each sample were kept the same and are outlined in table 3.24. The injection rate was set to 300 µl/hr.

Table 3.24: Instrument settings used for sample detection and analysis

Capillary	End Plate Offset	Skimmer	CapExit	Oct 1 DC	Oct 2 DC	Trap Drive	Oct RF	Lens 1	Lens 2
-4667	-500	44.8	158.3	5.38	2.00	50.0	291.7	-3.0	-53.5

In addition to the detection settings the ionisation settings were set as follows: Nebulizer 20 psi, Dry Gas Rate 8 L/min and the Dry Temperature set to 300 °C.

ESI-MS analysis of the *S. meliloti* 2011 siderophore extract

The *S. meliloti* 2011 extract was analysed by the method described previously. The key spectra of interest are of that obtained from the whole sample, the isolation of the molecular ion at 531, MS^2 analysis of the fragmentation of this ion and MS^3 analysis of the fragmentation of the main ion formed from the MS^2 analysis.

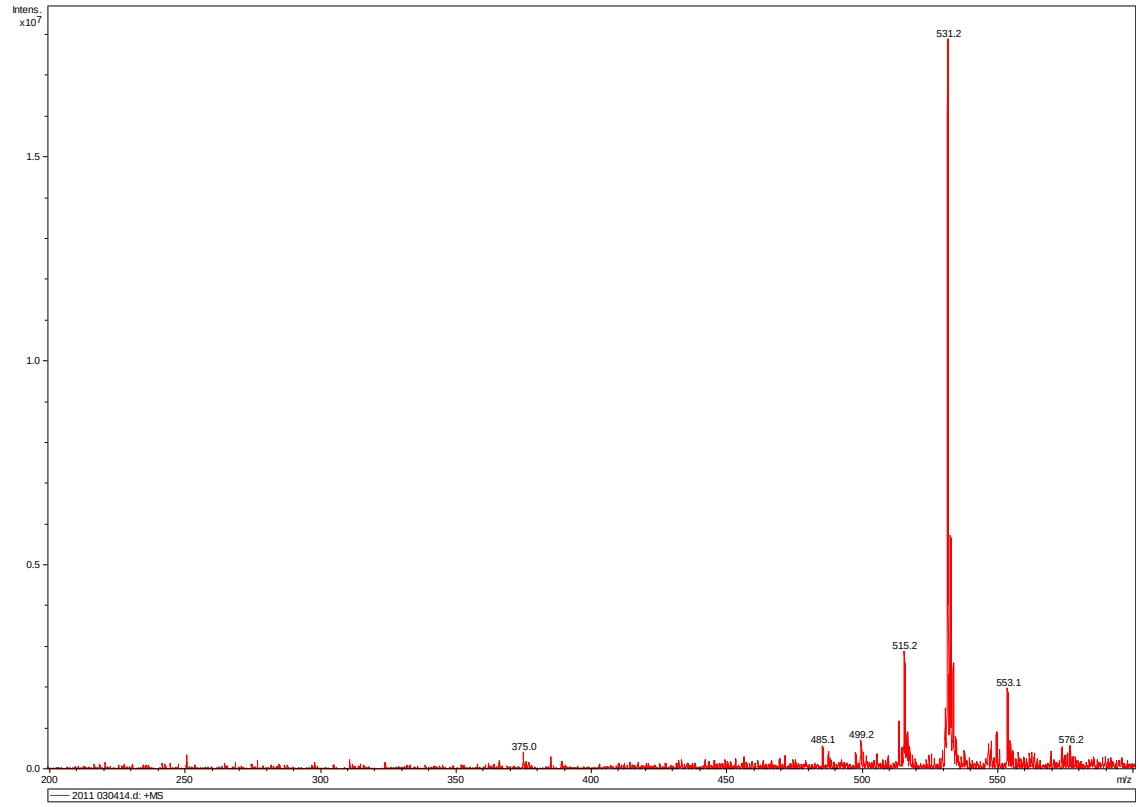


Figure 3.22: Spectrum of the whole extracted siderophore sample from *S. meliloti* 2011

The analysis of the whole siderophore extract from *S. meliloti* 2011 clearly shows the most prominent constituent of the sample to be of an m/z of 531.2 which is only 0.1 Da below the theoretical protonated molecular mass of 531.3 Da for rhizobactin 1021. The heavier ions observed could represent sodium adduct forms of molecular rhizobactin 1021 which is common in ESI-MS. The ion at 553.1 m/z constitutes a single sodium adduct form and the ion at 576.2 m/z would represent a disodium adduct form. Smaller ions may be explained by partial degradation of the siderophore or more simply may be contaminants from the extraction process. Notably there is no peak representative of the ferric form of the siderophore which would indicate that the extraction procedure results in a near complete deferration of the siderophore. Table 3.25 summarises the main features shown in this spectrum and gives possible explanations for their presence.

Table 3.25: Explanation of the prominent peaks from the analysis of the *S. meliloti* 2011

Molecular Formula	Theoretical molecular weight	Observed m/z
$[M+H]^+$	531.3 Da	531.2 Da
$[M+Na]^+$	553.3 Da	553.1 Da
$[M+2Na]^+$	576.3 Da	576.2 Da
$[M+H]^+$-[O]	515.3 Da	515.2 Da
$[M+H]^+$-[2O]	499.3 Da	499.2 Da
$[M+H]^+$-[2O]-[N]	485.3 Da	485.1 Da

M= Rhizobactin 1021 molecular weight, 530.3 Da

There is slight variation between the theoretical molecular weight and the observed molecular weight which may be as a result of the level of resolution provided by the ESI-MS instrument used. A higher resolution mass spectrometer would give a molecular weight to three decimal places and would negate any discrepancies introduced by rounding of the detected values. Regardless of this there is still a high degree of agreement between the theoretical molecular weight and the observed molecular weight with no more than a 0.2 Da discrepancy for any possibility.

The prominence of the 531.2 m/z peak will allow for a very clear fragmentation pattern to be ascertained. This ion was isolated by the mass spectrometer and subjected to fragmentation referred to as MS^2. Figure 3.23 is the spectrum observed for the fragmentation of the isolated 531.2 ion.

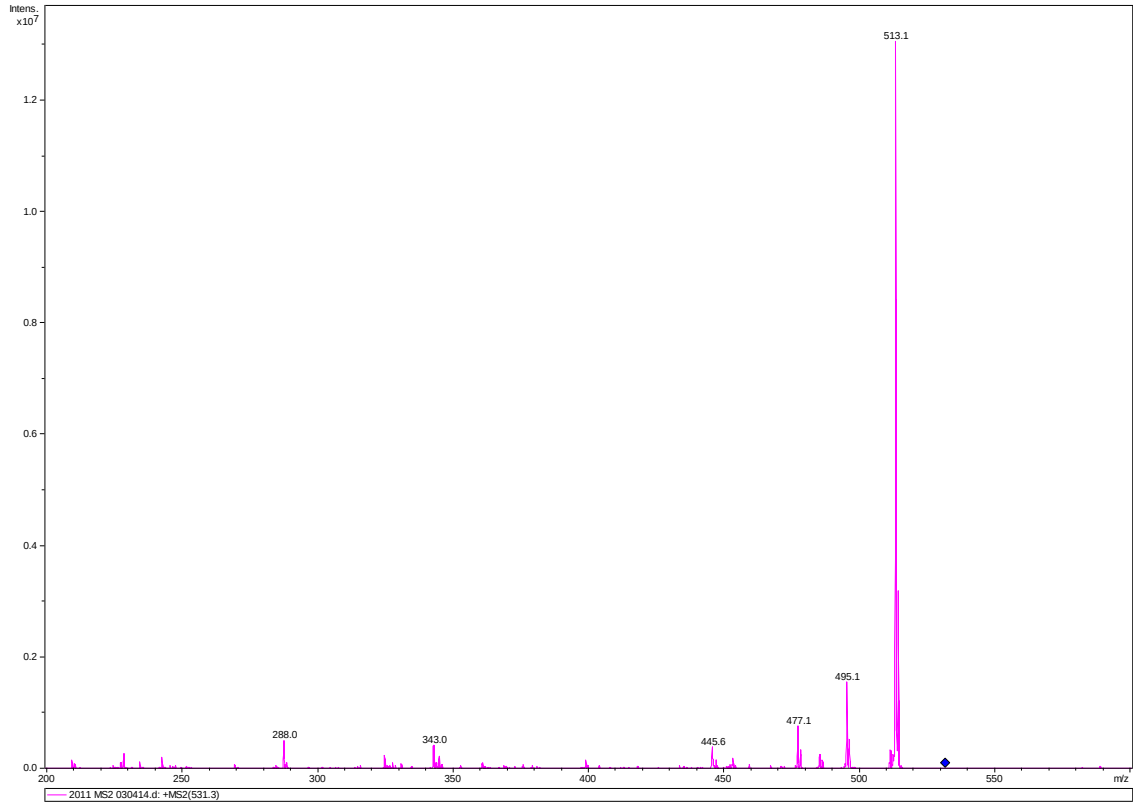


Figure 3.23: Spectrum of the isolated rhizobactin 1021 ion fragmentation pattern

As the 531.2 Da ion is isolated prior to fragmentation all background present in the previous spectrum is no longer observable. The position of the precursor ion is marked with a small blue diamond. The most prominent product ion resulting from the fragmentation has an m/z of 513. There are a number of clear peaks also identified and these are listed in table 3.26.

Table 3.26: Product ions from the fragmentation of the rhizobactin 1021 precursor ion

Observed m/z	Δ from 531.2	Possible molecular formula	Theoretical molecular weight
513.1 Da	18.1	$[M+H]^+$-$[H_2O]$	513.3 Da
495.1 Da	36.2	$[M+H]^+$-$2[H_2O]$	495.3 Da
477.1 Da	54.2	$[M+H]^+$-$3[H_2O]$	477.3 Da
445.6 Da	85.7	$[M+H]^+$-$3[H_2O]$-$[NO+H_2]$	445.3 Da
343.0 Da	188.3	Unknown	—
288.0 Da	243.3	Unknown	—

M= Rhizobactin 1021 molecular weight, 530.3 Da

The possible molecular formulae presented in table 3.26 are only best fit formulae, as loss of any combination of the elements found in rhizobactin 1021 adding up to the correct molecular weight could be correct. However the formulae given here are the most likely for the figures attained.

As observed previously there is no exact match between the theoretical molecular weight and the observed m/z in the spectrum. Each ion is consistently off by 0.3 Da, the reason for which is unclear however may be explained by the theoretical molecular weight being calculated using the relative atomic mass. The relative atomic mass is a weighted average of the atomic mass of all the known isotopes of a given element and can result in atomic mass values that are not whole numbers. If the theoretical molecular weight is recalculated using only the exact atomic mass of the most abundant isotope of the elements involved then this 0.3 Da discrepancy reduces to 0.1 Da for the three heaviest product ions but increases to 0.6 Da for the 445.6 Da daughter ion. However as there is a high level of agreement between all the values this slight variance is not of concern.

To obtain further data on the fragmentation of rhizobactin 1021 MS^3 analysis was carried out on the most prominent MS^2 product ion observed at 513.1 m/z. As with the previous fragmentation the 513.1 m/z ion was isolated prior to fragmentation therefore removing all other product ions from the MS^2.

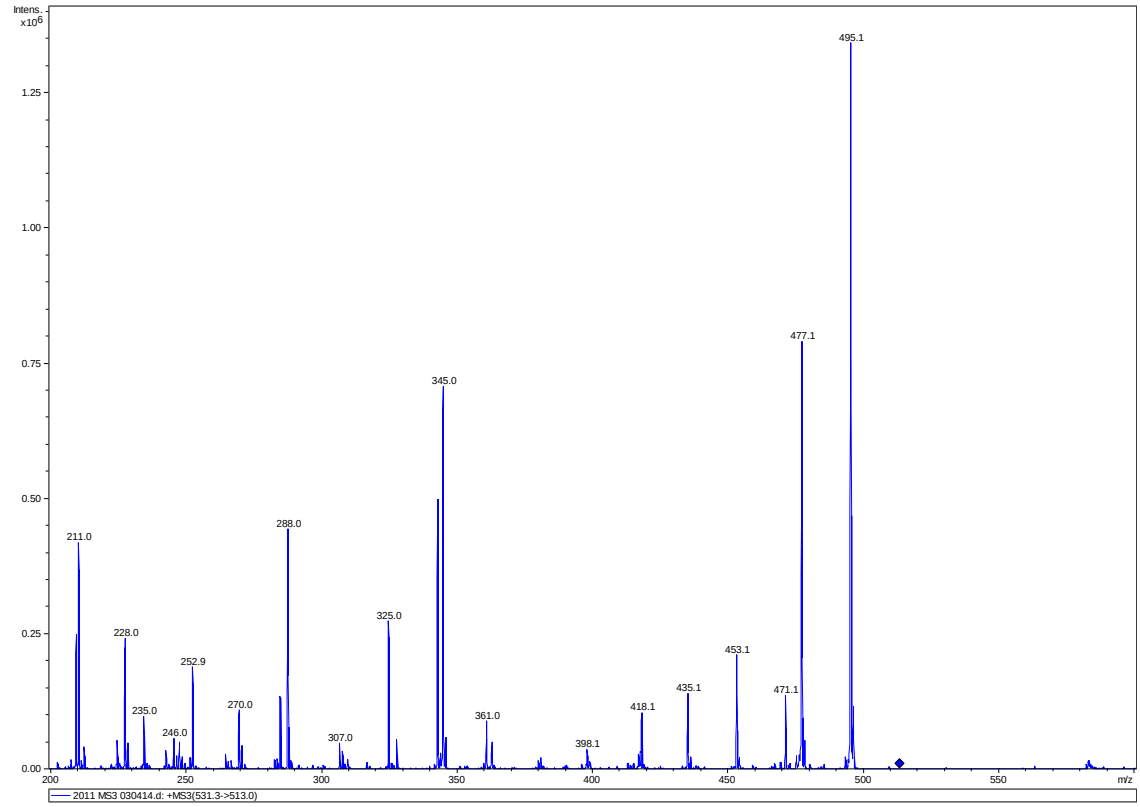


Figure 3.24: Spectrum of the MS^3 analysis of the rhizobactin 1021 precursor ion

The spectrum for the fragmentation of the 513.1 ion is far more complex than the spectrum attained from the fragmentation of the rhizobactin 1021 precursor ion. Due to the complexity of the spectrum it would be difficult to determine the degradation that resulted in the occurrence of each peak. However for the heavier peaks a plausible pattern of degradation of the ion can be elucidated. Table 3.27 describes the possible element loss to achieve the observed m/z.

Table 3.27: Product ions from MS^3 of the rhizobactin 1021 precursor ion

Observed m/z	Δ from 513.1	Possible molecular formula	Theoretical molecular weight
495.1 Da	18	$[P+H]^+$-$[H_2O]$	495.1 Da
477.1 Da	36	$[P+H]^+$-2$[H_2O]$	477.1 Da
471.1 Da	42	$[P+H]^+$-$[CH_2C=O]$	471.1 Da
453.1 Da	60	$[P+H]^+$-$[CH_2C=O]$-$[H_2O]$	453.1 Da
435.1 Da	78	$[P+H]^+$-$[CH_2C=O]$-2$[H_2O]$	435.1 Da
418.1 Da	95	$[P+H]^+$-$[NH_2C_2H_3O]$-2$[H_2O]$	418.1 Da
398.1 Da	115	$[P+H]^+$-$[NH_2C_2H_5O]$-3$[H_2O]$	398.1 Da

P: Represents a non-protonated 513.1 ion

As can be seen the possible molecular formulae match both the observed m/z and the theoretical molecular weight exactly. As before the possible molecular formulae are only proposals that are best fit to the observed change in m/z value.

This analysis of rhizobactin 1021 has allowed for clear fragmentation patterns to be determined that will serve as a template to analyse each of the extracts from *S. meliloti* 2011*sma2339*, *S. meliloti* 2011*rhbD* (pCC103) and *S. meliloti* 2011*rhbA62*.

ESI-MS analysis of the *S. meliloti* 2011*sma2339* siderophore extract

The siderophore extract from *S. meliloti* 2011*sma2339* was analysed using the identical settings used to analyse the *S. meliloti* 2011 siderophore extract. Throughout these analysis key indicators for the presence of rhizobactin 1021 was a priority.

As carried out with the *S. meliloti* 2011 extract the entire sample was first analysed to determine both the abundance of rhizobactin 1021 and to identify possible unusual ions. Figure 3.25 is the spectrum obtained from the whole sample extract from *S. meliloti* 2011*sma2339*.

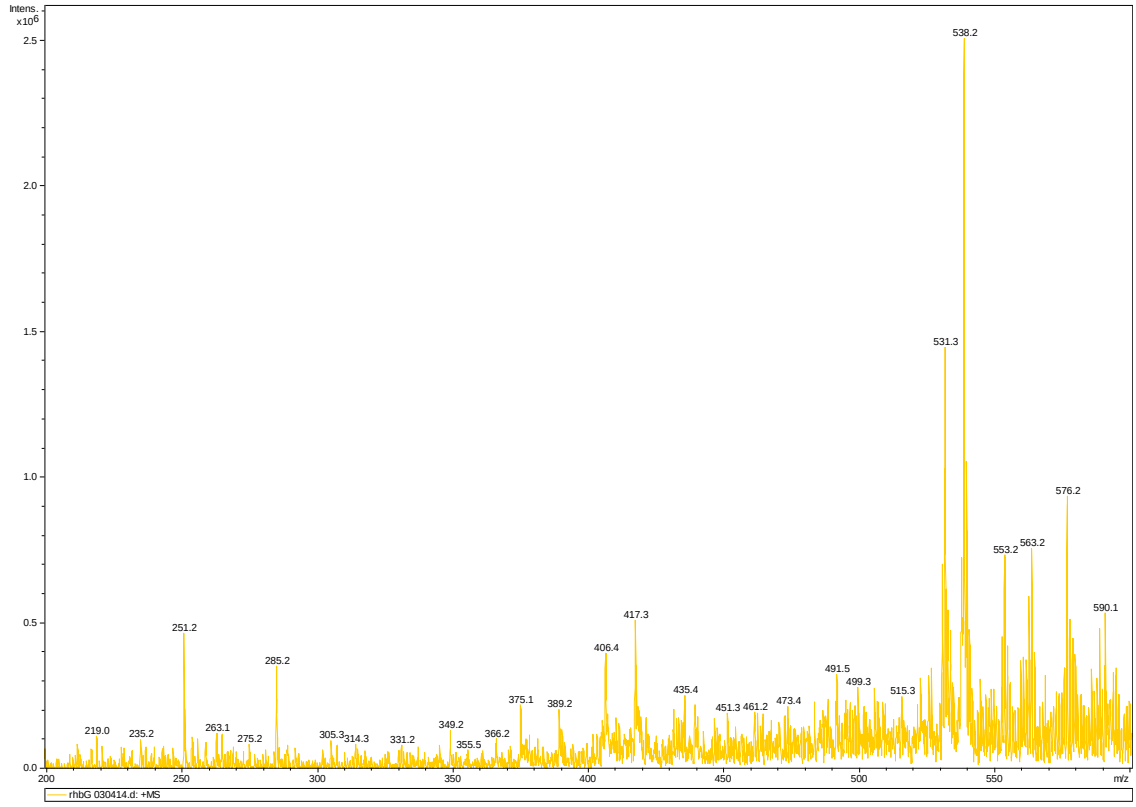


Figure 3.25: Spectrum of the whole extracted siderophore sample from *S. meliloti* 2011*sma2339*

There is a clear peak at 531.3 m/z representing the protonated rhizobactin 1021 molecular ion. In addition a number of definable peaks are present that match the rhizobactin 1021 forms found in the *S. meliloti* 2011 sample. These matches are outlined in table 3.28.

Table 3.28: Rhizobactin 1021 related m/z values matching values observed in *S. meliloti* 2011

Molecular Formula	Observed m/z for *S. meliloti* 2011	Observed m/z for *S. meliloti* 2011*sma2339*
$[M+H]^+$	531.2 Da	531.3 Da
$[M+Na]^+$	553.3 Da	553.2 Da
$[M+2Na]^+$	576.3 Da	576.2 Da
$[M+H]^+$-[O]	515.3 Da	515.3 Da
$[M+H]^+$-[2O]	499.3 Da	499.3 Da
$[M+H]^+$-[2O]-[N]	485.3 Da	Unclear

M= Rhizobactin 1021 molecular weight, 530.3 Da

As there is such a close match to the rhizobactin 1021 related m/z values in both samples it is probable that rhizobactin 1021 is present in the sample. As expected the abundance of the 531.3 ion is far below that observed in the *S. meliloti* 2011 sample which reflects the lower levels of production observed by the extraction by the CAS assay. There is also no clear peak representing a ferrated siderophore indicating that the

siderophore in this sample is deferrated similar to the *S. meliloti* 2011 sample. One striking difference between the *S. meliloti* 2011 and *S. meliloti* 2011*sma2339* is the presence of a prominent peak at 538.2 m/z in the *S. meliloti* 2011*sma2339* sample. It is unlikely that this ion is related to rhizobactin 1021 however the ion was isolated and fragmented to confirm that its fragmentation pattern was not similar to rhizobactin 1021. The spectrum obtained from this fragmentation is given in figure 3.26.

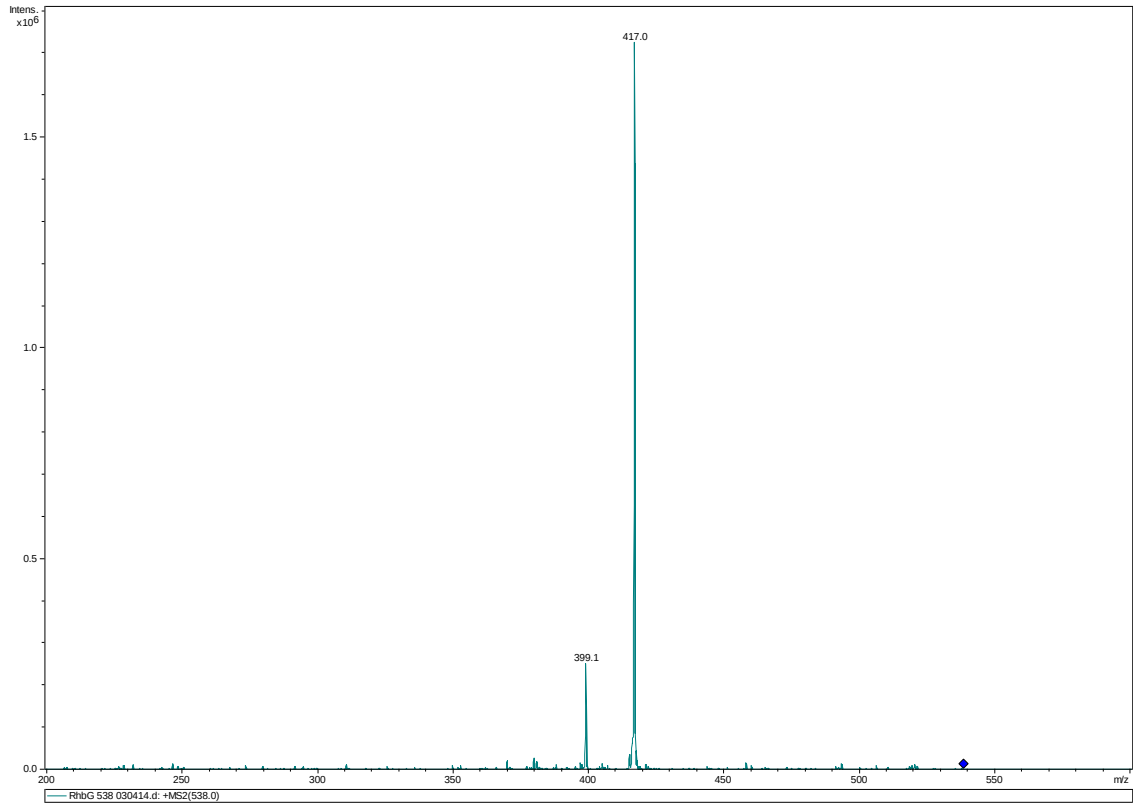


Figure 3.26: Spectrum of the fragmentation pattern observed from the 538.2 ion in the *S. meliloti* 2011*sma2339* extract

The fragmentation of 538.2 ion yields two prominent product ions with m/z values of 417.0 and 399.1. As neither of these values are representative of any product ion in the rhizobactin 1021 MS^2 or MS^3 analysis, the 538.2 ion was classed as a contaminant in the sample and is not of interest.

The 531.3 m/z ion was subjected to fragmentation to confirm its identity as rhizobactin 1021. If a highly similar fragmentation pattern is observed in both the MS^2 and MS^3 analysis this will be considered to be confirmation that it is rhizobactin 1021. The spectrum of the MS^2 analysis is given in figure 3.27.

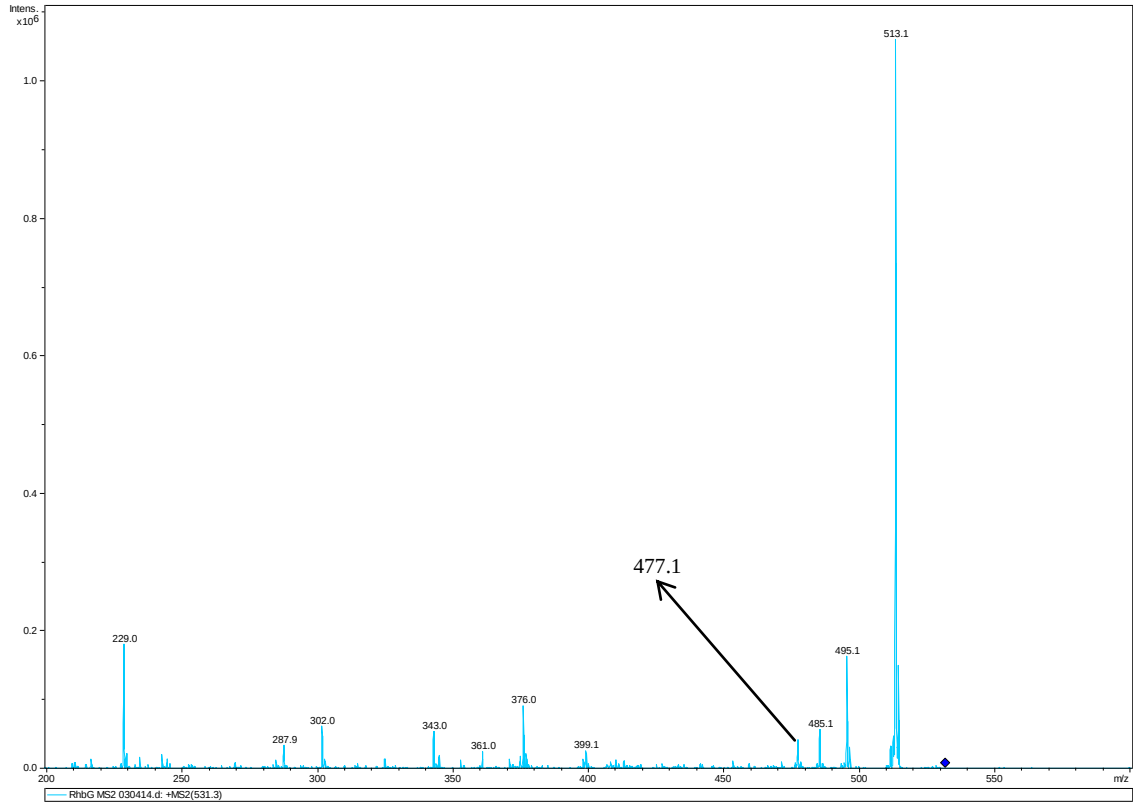


Figure 3.27: Spectrum of the isolated 531.3 m/z ion fragmentation pattern from the *S. meliloti* 2011*sma2339* extract

There are five clear matching peaks that align with the MS^2 data obtained from the *S. meliloti* 2011 extract. These are given in table 3.29.

Table 3.29: MS^2 product ions from *S. meliloti 2011sma2339* matching *S. meliloti* 2011

Molecular Formula	Observed m/z for *S. meliloti* 2011	Observed m/z for *S. meliloti* 2011*sma2339*
$[M+H]^+$-$[H_2O]$	513.1 Da	513.1 Da
$[M+H]^+$-2$[H_2O]$	495.1 Da	495.1 Da
$[M+H]^+$-3$[H_2O]$	477.1 Da	477.1 Da
Unknown	343.0 Da	343.0 Da
Unknown	288.0 Da	287.9 Da

M= Rhizobactin 1021 molecular weight, 530.3 Da

In addition to these matching peaks there are six peaks that do not match the rhizobactin 1021 fragmentation profile. These may be explained by the higher background level in the *S. meliloti* 2011*sma2339* sample with the extra peaks being formed by fragmentation of co-isolated ions. In order to isolate an ion, a bandwidth must be set to allow for sufficient ions to be collected and in the case of these samples the bandwidth was set to four. A lower bandwidth led to a large reduction of the isolated peak height that could not be further fragmented. In practical terms this bandwidth results in all ions

from m/z 529.3-533.3 being isolated and fragmented. This was not a problem when analysing the *S. meliloti* 2011 sample as the 531 peak was far more intense than the background noise however there are a number of lesser peaks either side of the 531 m/z peak in the *S. meliloti* 2011*sma2339* sample. This is the most likely source of the contaminants in the MS^2 spectrum.

To completely confirm the identity of the 531 ion in the *S. meliloti* 2011*sma2339* sample the MS^3 fragmentation was carried out on the most prominent 513.1 m/z ion. The spectrum from this fragmentation is given in figure 3.28.

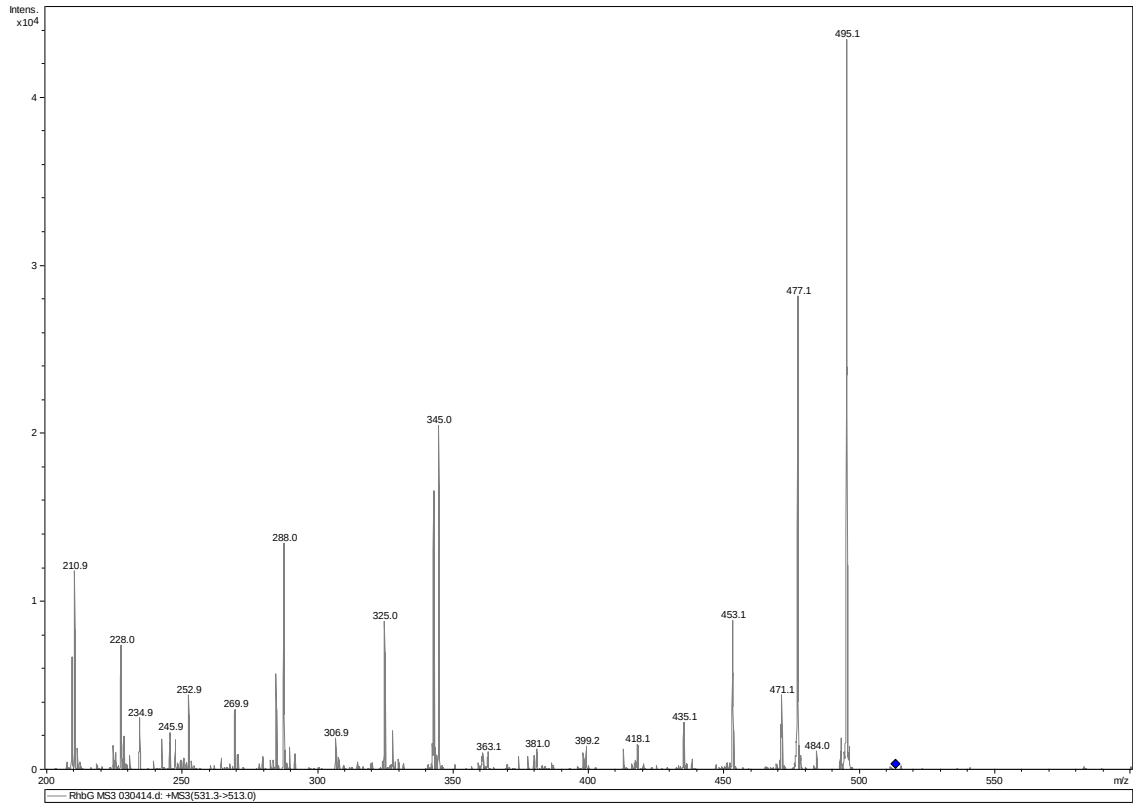


Figure 3.28: Spectrum of the isolated 513.1 m/z ion fragmentation pattern from *S. meliloti* 2011*sma2339* extract

Analysis of the MS^3 data obtained from the *S. meliloti* 2011*sma2339* extract overlays almost exactly with the MS^3 data from the *S. meliloti* 2011 extract. There are only three peaks in the *S. meliloti* 2011*sma2339* extract that are not present in the *S. meliloti* 2011 sample namely 484 m/z, 381 m/z and 363.1 m/z. Closer observation of the spectrum in figure 3.28 shows a number of peaks at these points that are not annotated. Annotation is dictated by the mass spectrometers software and these peaks may not have been prominent at the time of imaging. Even with these slight discrepancies it is clear that the fragmentation is from the same structured molecule.

Analyses of the spectra obtained from the *S. meliloti* 2011*sma2339* extract clearly show that there is intact rhizobactin 1021 present in the sample. This is significant as it shows that the Challis model for the biosynthesis of rhizobactin 1021 given in figure 3.3 B is incorrect as the Sma2339 protein does not add the lipid to the siderophore. Also there is no indication of a non lipidated siderophore in the spectrum of the whole extract as rhizobactin 1021 without a lipid moiety would closely resemble the siderophore schizokinen and show a molecular weight of 421 Da.

As the *S. meliloti* 2011*sma2339* mutant produces rhizobactin 1021, albeit at a reduced rate, this shows that Sma2339 does not fulfil a critical single role in the biosynthesis of rhizobactin 1021 with regard to structure. The reduction in production does prove that it is involved but its role in assembling rhizobactin 1021 can be sufficiently complemented, most likely by RhbD, as shown by figure 3.17. This result will be discussed further in the Conclusions section with regard to unpublished results obtained as part of a study into the interactions between each of the rhizobactin 1021 biosynthesis proteins. This analysis demonstrated that Sma2339 shows more interaction than any of the other biosynthesis proteins, indicating that it may by playing a structural role, localising the biosynthesis proteins together.

ESI-MS analysis of the *S. meliloti* 2011*rhbD* (pCC103) siderophore extract

As the absence of the Sma2339 protein has been shown to have no effect on the structure of the produced siderophore, the structure of the *S. meliloti* 2011*rhbD* (pCC103) siderophore was also determined. The siderophore extract was analysed by the same process as the *S. meliloti* 2011 and *S. meliloti* 2011*sma2339* extracts. The spectrum from this analysis is given in figure 3.29.

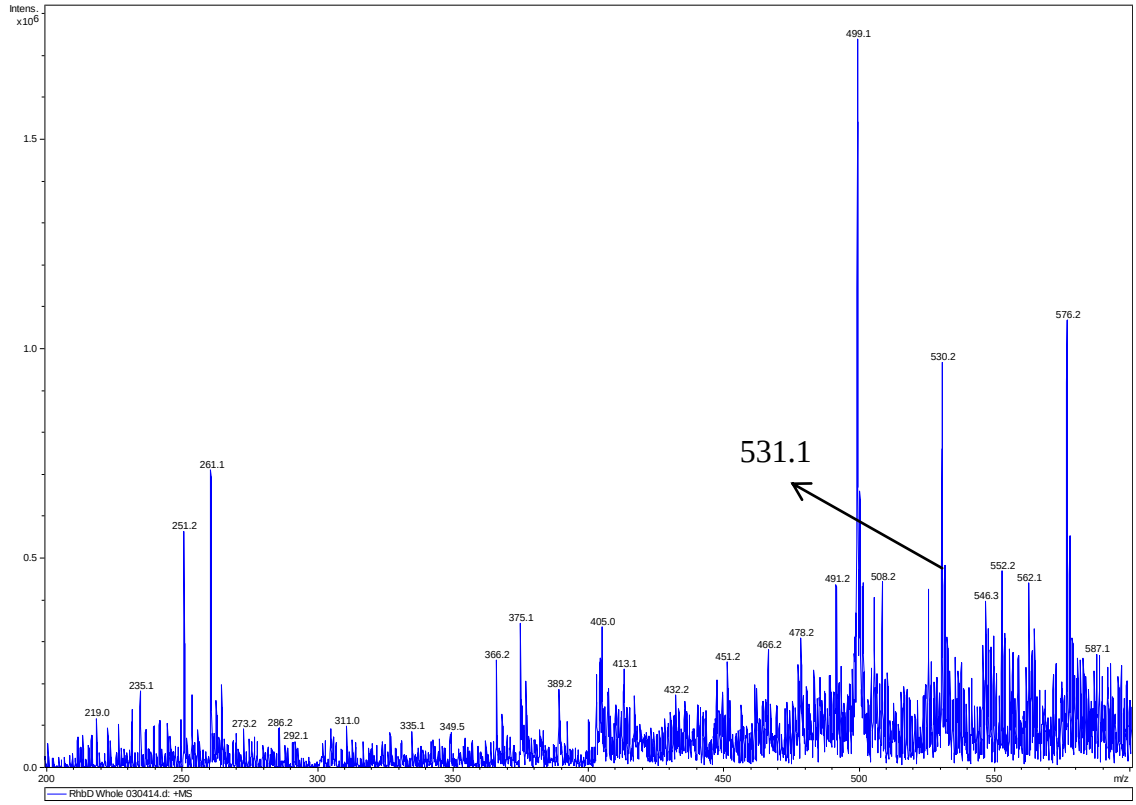


Figure 3.29: Spectrum of the whole extracted siderophore sample from S. meliloti 2011*rhbD* (pCC103)

As can be seen in figure 3.29 there is no distinct peak at 531.3 representing the protonated rhizobactin 1021 molecule. Zooming in on the region close to the 531.3 m/z position reveals a peak at 531.1 m/z. Identification of possible sodium adducts at positions 553.3 m/z and 576.3 m/z shows there to be a clear peak for the disodium adduct with an observed value of 576.2 m/z with no identifiable peak at the 553.3 m/z size. In addition to the sodium adduct peaks the peak for the possible dideoxyrhizobactin 1021 at 499.1 is most prominent in the sample. This is interesting as it may suggest that an aberrant siderophore is being produced due to the lack of the primary acetyltransferase. This peak is also observed in the first two samples however it is of far less abundance in relation to a 531.3 m/z peak. It is likely that removing two

oxygen molecules from rhizobactin 1021 would abolish or greatly reduce its ability to chelate iron as all the oxygen molecules are predicted to play a role in chelation. There is also a prominent contaminant peak at position 530.2 m/z which would represent the non-protonated weight of rhizobactin 1021. However this is seen as a contaminant as it was not observed in either of the two previous samples. Indications for a possible 420 m/z value representative of a structure lacking a lipid moiety were also not found.

In summary of the points in relation to this spectrum it is clear that due to the low levels of siderophore in the sample it is difficult to definitively state if rhizobactin 1021 is present or absent in the sample. There are indications for its presence but these are not as clear as observed in the *S. meliloti* 2011*sma2339* siderophore extract.

The relationship between the 499 m/z ion and rhizobactin 1021 could not be confirmed. This would require a standard fragmentation pattern of rhizobactin 1021 missing two oxygens to be obtained. This molecule would have to be synthesised as *S. meliloti* 2011 does not produce it in vast quantities. The peak at 531.1 m/z was subjected to MS^2 analysis to determine if it is rhizobactin 1021 the spectrum from which is given below.

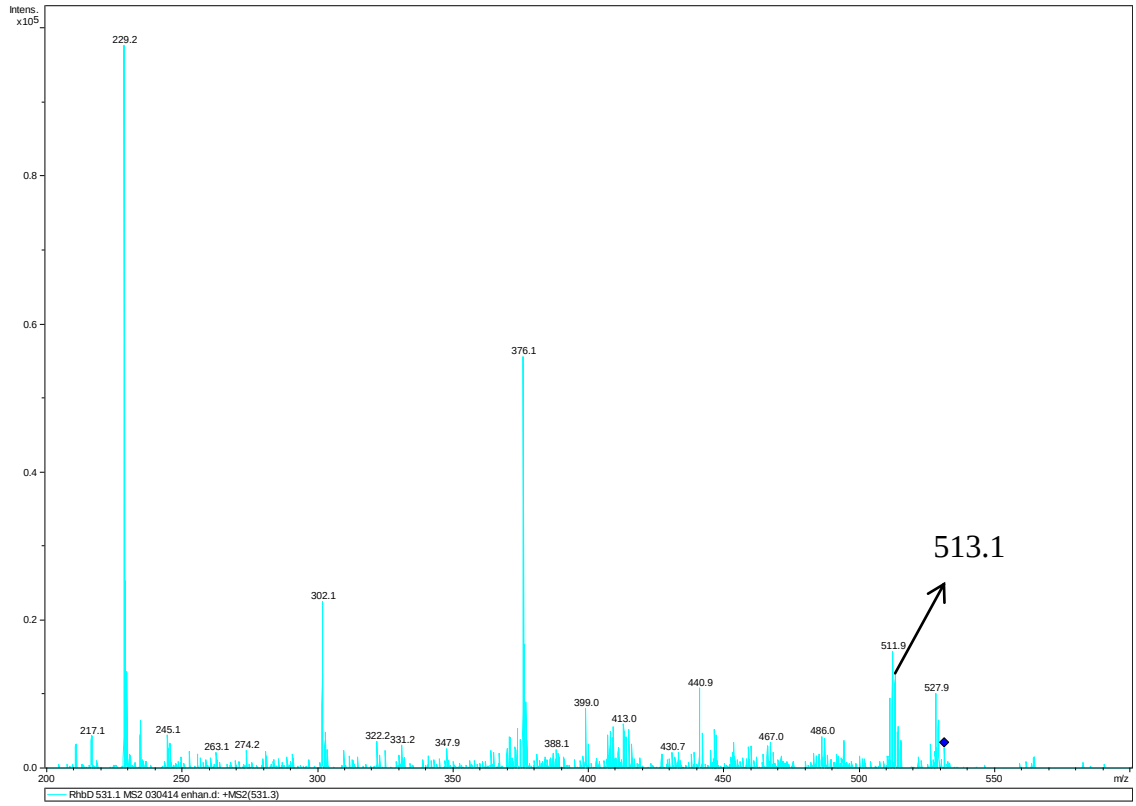


Figure 3.30: Spectrum of the isolated 531.1 m/z ion fragmentation pattern from the S. meliloti 2011*rhbD* (pCC103) extract

Comparing figure 3.30 to the MS^2 data in figure 3.23 there is very little similarity. The peak 513.1 m/z is not prominent but is present with the other matches to the rhizobactin 1021 MS^2 spectrum being 399.1 m/z, 376.1 m/z and 229.2 m/z. The 399.1 m/z and 376.1 m/z peaks are only visible on a zoomed in version of the rhizobactin MS^2 spectrum and cannot be seen in figure 3.23. However the peaks at position 376.1 m/z and 229.1 m/z can be ruled out as product ions formed from fragmentation of a contaminant, possible at 530.2 m/z. Considering figure 3.27 showing the fragmentation spectrum of the 531.3 m/z ion from the *S. meliloti* 2011*sma2339* extract along with figure 3.26 it is apparent that the peaks 376 m/z and 229 m/z are markedly larger in relation to the other peaks in the *S. meliloti* 2011*sma2339* sample. This would correlate with the greater peak height of contamination peaks close to 531.3 m/z in relation to the actual 531.3 m/z ion as can be seen in figure 3.25. Keeping this in mind when analysing the spectrum in figure 3.29, the whole extract from *S. meliloti* 2011*rhbD* (pCC103), the contaminants close to 531.3 m/z are more prominent than the 531.3 m/z ion peak. This should result in the peaks for 376 m/z and 299 m/z getting progressively larger in relation to the other peaks in the MS^2 spectra moving from *S. meliloti* 2011 through *S. meliloti* 2011*sma2339* to *S. meliloti* 2011*rhbD* (pCC103) which is the observed case. Further evidence for these being contaminants is shown by observation of the 302 m/z ion which is not present in *S. meliloti* 2011 but is present in the other samples and is more prominent in the *S. meliloti* 2011*rhbD* (pCC103) extract. Disregarding these contaminants there are only two matches between the MS^2 of 531.3 for the *S. meliloti* 2011 and *S .meliloti* 2011*rhbD* (pCC103) at positions 399 m/z and 513 m/z both of which are not prominent. As a result of the low intensity levels observed it is unlikely that an informative MS^3 spectrum can be achieved for the *S. meliloti* 2011*rhbD* (pCC103) extract but it was attempted for completeness. The observed spectrum for this MS^3 fragmentation is given in figure 3.31.

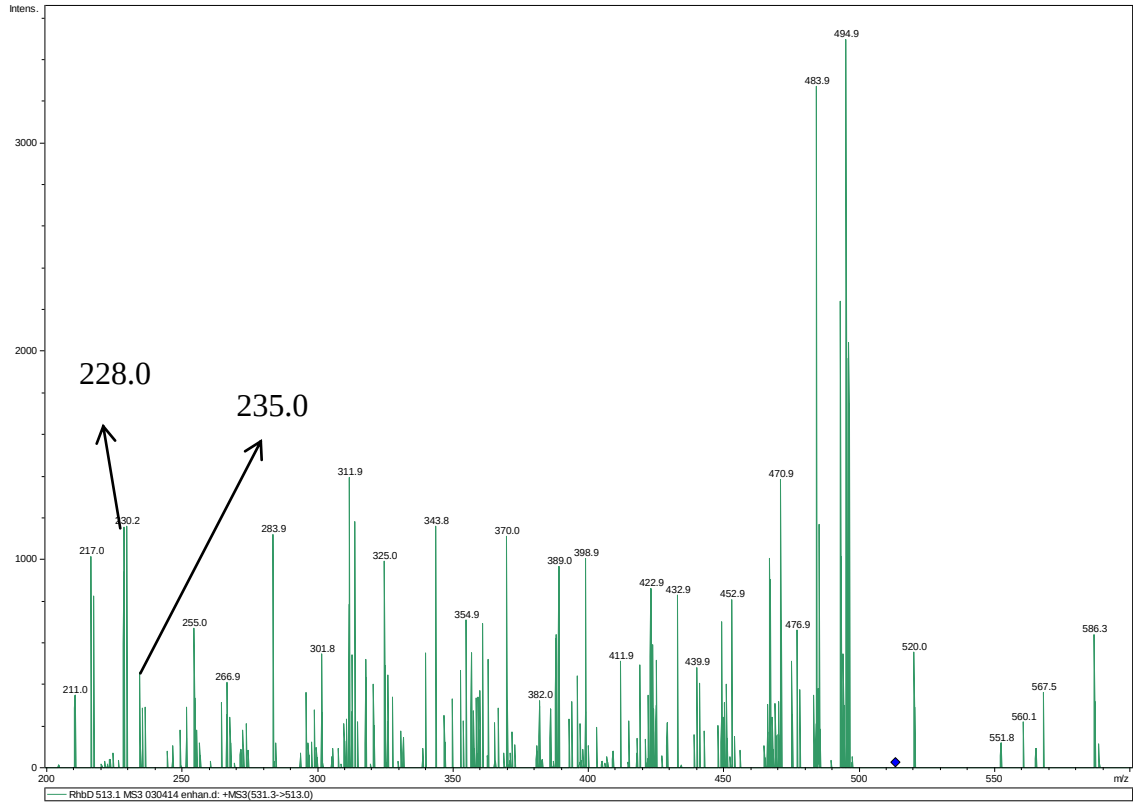


Figure 3.31: Spectrum of the isolated 513.1 m/z ion fragmentation pattern from *S. meliloti* 2011*rhbD* (pCC103) extract

Initial analysis of figure 3.31 suggests that there is no similarity to the MS^3 data obtained from the *S. meliloti* 2011 extract. However there are 11 matches in this spectrum to prominent peaks observed in the MS^3 analysis of rhizobactin 1021. These matches are outlined in table 3.30.

Table 3.30: Observed matches between MS^3 spectra for *S. meliloti* 2011 and *S. meliloti* 2011*rhbD* (pCC103)

***S. meliloti* 2011**	***S. meliloti* 2011*rhbD* (pCC103)**
495.1 Da	494.9 Da
477.1 Da	476.9 Da
471.1 Da	470.9 Da
453.1 Da	452.9 Da
398.1 Da	398.9 Da
345.0 Da	343.8 Da
325.0 Da	325.0 Da
270.0 Da	266.9 Da
235.0 Da	235.0 Da
228.0 Da	228.0 Da
211.0 Da	211.0 Da

This amount of matches strongly suggest that rhizobactin 1021 is indeed present in the *S. meliloti* 2011*rhbD* (pCC103) siderophore extract considering that there was only a

maximum number of 17 possible matches. This result is not as confident as the result for the *S. meliloti* 2011*sma2339* result due to two main factors: the very low intensity levels and the level of contamination. In figure 3.31 and with all the spectra, the intensity of the signal is scaled on the y axis. The intensity in figure 3.31 is ~10 fold less than that observed in the equivalent MS^3 sample for *S. meliloti* 2011*sma2339* and ~1000 fold for the MS^3 on the *S. meliloti* 2011 sample. This low level of intensity is near the limit of detection for the instrument and a lot of background is present. This is exemplified by the presence of ions between 520 m/z and 586 m/z which should not be present as the 513.1±2 ion was isolated prior to fragmentation and only lighter molecules should be present. Another factor occluding the result is the high level of contaminant ions very close to 513.1 m/z which are more prominent than the desired ion and will be co-isolated due to their proximity. These ions will also get fragmented and their product ions are present throughout the sample.

The results presented for *S. meliloti* 2011*rhbD* (pCC103) do not absolutely confirm the presence of rhizobactin 1021 but do give a very strong indication of its presence. If these indications are correct it has a number of implications for the biosynthesis of rhizobactin 1021. These will be discussed in detail at the end of the section but in brief it indicates that there must be an eighth protein closely involved in the production of rhizobactin 1021. This protein must also be a member of a separate cluster of genes as there are no other possible biosynthesis genes located proximal to the identified biosynthesis cluster.

ESI-MS analysis of the *S. meliloti* 2011*rhbA62* siderophore extract

The extract from the *S. meliloti* 2011*rhbA62* supernatant will provide an excellent negative control for the ESI-MS analyses presented above. Theoretically there should be very few similarities between the MS^2 and MS^3 data from the negative compared with the MS^2 and MS^3 data from the positive and mutant strains. The spectrum from the whole siderophore extract of the mutant is given the figure 3.32.

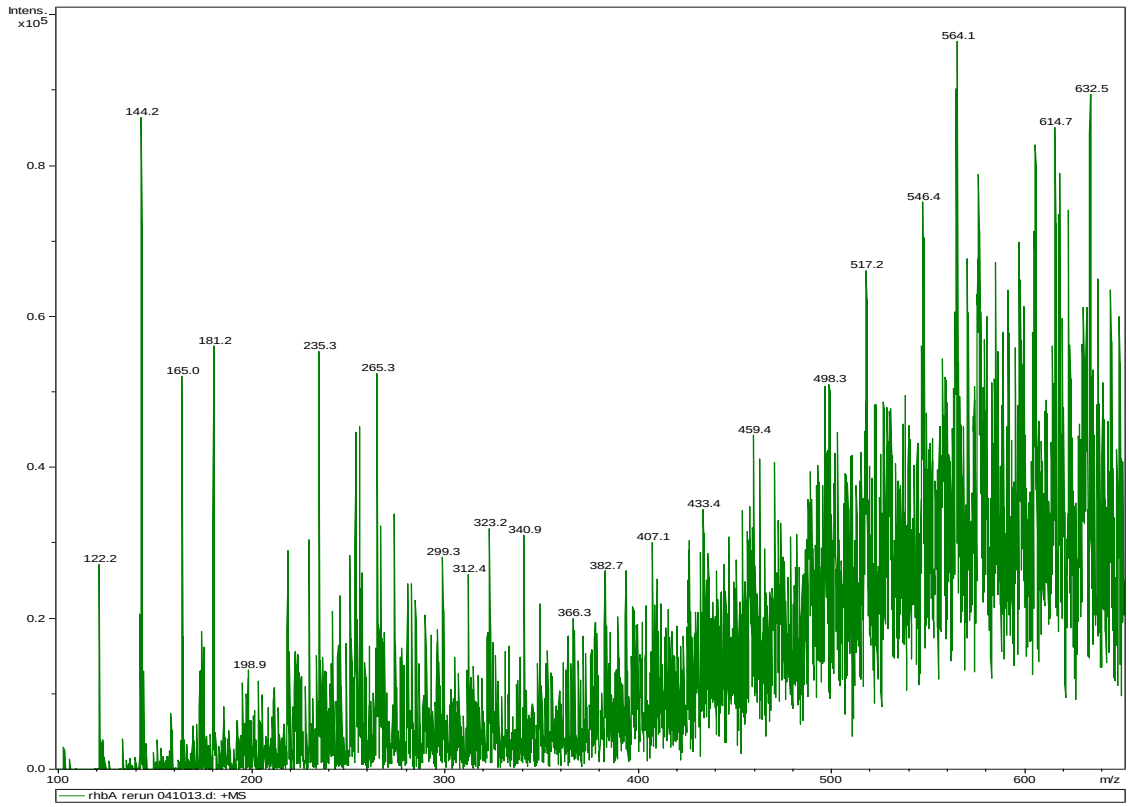


Figure 3.32: Spectrum of the whole extracted siderophore sample from *S. meliloti* 2011*rhbA62*

The whole sample spectrum for *S. meliloti* 2011*rhbA62* from figure 3.32 shows none of the prominent peaks that are associated with rhizobactin 1021. The main indicators would have been peaks representing the protonated siderophore at 531.3 m/z, peaks representing sodium adducts at 553 m/z and 576 m/z, and peaks representing possible degradation products at 515 m/z and 499 m/z, none of which are present in high levels. Regardless of not being able to see clear peaks the analysis was continued as before to give a negative for each spectrum. An isolation of ions around position 531.3 m/z was carried out and these were fragmented as before. The isolation procedure isolated a main peak at 529.8 m/z with possible peaks close to 531 m/z. The fragmentation spectrum of the isolated ions is given in figure 3.33.

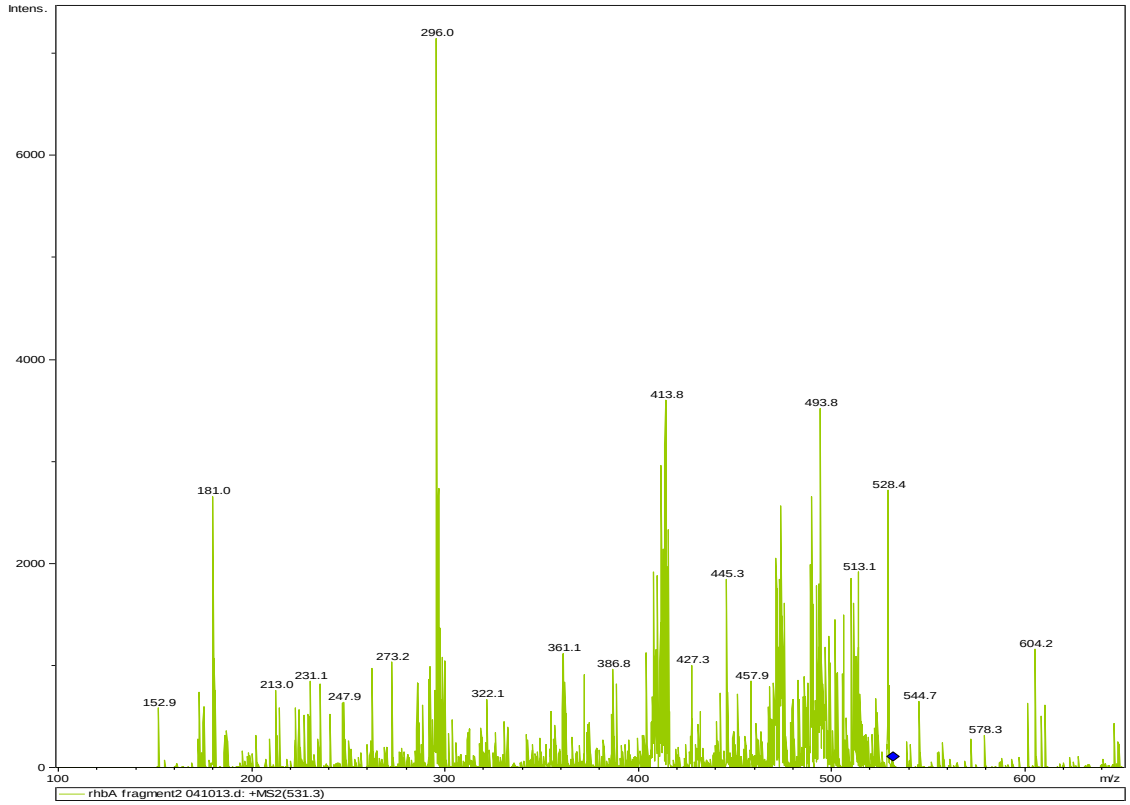


Figure 3.33: Spectrum of the isolated 531.3 m/z ion fragmentation pattern from the *S. meliloti* 2011*rhbA62* extract

There are three matches between the MS^2 for rhizobactin 1021 and the spectrum in figure 3.33 at positions 513.1 m/z, 445.1 m/z and 361.1 m/z. As observed with the MS^3 spectrum from the *S. meliloti* 2011*rhbD* (pCC103) the intensity levels are low indicating that there was very little of the parent ions to work with.

The 513.1 m/z ion was isolated and fragmented to ascertain if it is a rhizobactin 1021 related ion or a coincidental ion at that size. Figure 3.34 is the spectrum obtained from the MS^3 analysis.

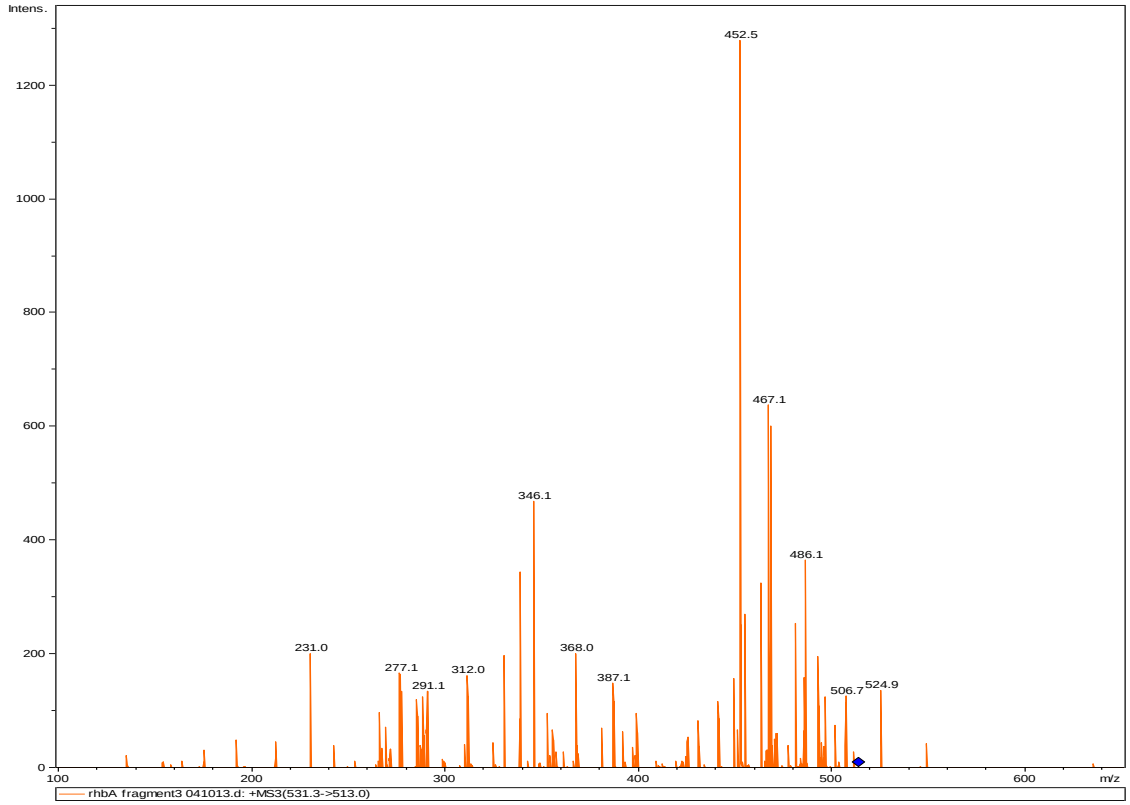


Figure 3.34: Spectrum of the isolated 513.1 m/z ion fragmentation pattern from *S. meliloti* 2011*rhbA62* extract

As carried out previously the 513.1 m/z ion was isolated, fragmented and the fragmentation pattern recorded. As expected for a negative control the fragmentation pattern does not resemble the MS^3 pattern obtained for rhizobactin 1021. There is no overlap between the peaks obtained in the rhizobactin 1021 analysis and the peaks in figure 3.34. These results provided a good baseline to judge the possible presence of rhizobactin 1021 as the MS^2 data in figure 3.33 did show three peaks that were associated with the presence of rhizobactin 1021 without there being siderophore present.

3.10 Summary and Discussion

The data presented here taken in their entirety creates an interesting argument for the role of the two acetyltransferases RhbD and Sma2339 in the biosynthesis of rhizobactin 1021. The Challis model for rhizobactin 1021 biosynthesis postulates that each of the proteins fulfils separate roles that are somewhat similar but are non-overlapping i.e. they both acetylate the same 1,3 diaminopropane derived substrate but with different acetyl length groups. The Lynch model proposes the same role for RhbD as the Challis model but makes no mention of the Sma2339 protein other than to state its presence and suggests an unnamed protein in the role of lipid addition. If the more complete proposed model, the Challis model, was correct then mutation of the *sma2339* gene would have resulted in the production of schizokinen in place of rhizobactin 1021. This was not the observed case, instead Sma2339 appears to be a redundant biosynthesis protein with regard to siderophore structure in that its abolition does not interfere with the final structure of the produced siderophore. Neither of the proposed models is fundamentally incorrect, the results presented here suggest a more subtle role for the two proteins in rhizobactin 1021 synthesis relative to one another.

The initial in silico comparison of the rhizobactin 1021 operon with biosynthesis operons of similarly structured siderophores provided the first evidence that perhaps Sma2339 was not fulfilling the proposed function of lipid addition. This analysis showed that regardless of the terminal acetyl group on the siderophore there was no apparent correlation between acetyl chain length with the size or number of acetyltransferase proteins associated with each cluster. Perhaps the most striking example of this is the close relationship between the biosynthesis operons of aerobactin and of the ochrobactins A-C. Each siderophore is assembled through the activity of four biosynthesis proteins homologous to RhbCDEF of the rhizobactin 1021 operon but in the case of aerobactin only a one carbon chain acetyl group is added to the siderophore which contrasts greatly with the variety of carbon chain length acetyl groups added by the *O. anthropi* species. This partial result taken alone indicates that the final siderophore structure may be dictated by functions not located as part of the core siderophore biosynthesis cluster. In addition to this the same conclusions can be drawn from the biosynthesis clusters of the 1,3 diaminopropane based siderophores. The biosynthesis clusters for schizokinen, acinetoferrin, rhizobactin 1021 and synechobactin A-C all comprise a core gene operon of six biosynthesis genes homologous to

rhbABCDEF. There are a number of differences between each operon in that the acinetoferrin operon appears to be split between two locations, the schizokinen *rhbF* homolog is not proximal to the other five biosynthesis genes, the *rhbF* homolog for the synechobactins biosynthesis is also not proximal to the other biosynthesis genes and has a C-terminal acetyltransferase domain and there being two acetyltransferases associated with rhizobactin 1021. Whereas these are notable differences in biosynthesis cluster arrangement, none correlate with acetyl chain length on the siderophore as schizokinen has a basic one carbon chain acetyl group, acinetoferrin is doubly acetylated with octenoic acid, the synechobactins are singularly acetylated with either octenoic acid, decenoic acid or dodecanoic acid, and rhizobactin 1021 being singularly acetylated with a decenoic acid moiety. A relationship between acetyltransferases and acetyl group length could not be found by BLASTp analysis or by protein alignment by the alignment tool MUSCLE either. This dataset resulted in the role of both the acetyltransferases, RhbD and Sma2339 being assessed functionally as neither could have its exact role elucidated by bioinformatics alone.

To assess the role of the two acetyltransferases three knockout mutants were created. Two of these mutants were single gene knockouts of each gene with the third mutant being a double mutant. To counter the polar effect observed in the *rhbD* mutant strains, the genotype had to be complemented with *rhbEF in trans* to illicit siderophore production. Siderophore production assessment of these mutants concluded that each of the single gene mutants were capable of producing siderophore although the production levels were drastically decreased from the levels observed in the *S. meliloti* 2011 parental strain. The double mutant was defective in siderophore production which confirms the key role acetyltransferases play in rhizobactin 1021 production. These results were obtained through the use of both the CAS plate assay and the CAS liquid assay. The CAS plate assay allowed for an initial screen for siderophore production to be carried out which gave a clear phenotype for each strain. The liquid CAS assay allowed the relative production levels of each of the positive producers to be assessed and showed a ~10 fold reduction in production in the absence of *sma2339* and a ~200 fold reduction in the absence of *rhbD*. These results indicate that each of the acetyltransferases can complement one another, however RhbD is the primary acetyltransferase as the greater drop in production was observed in its absence. Having

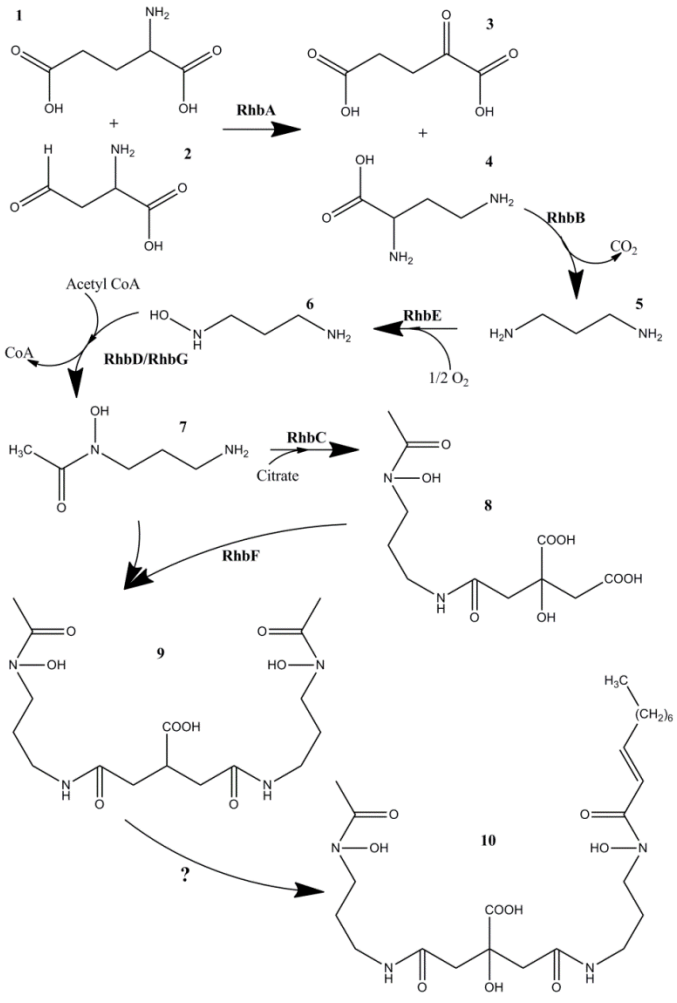
determined the production phenotype for each strain the siderophore structure was investigated by ESI-MS.

The two control sample extracts *S. meliloti* 2011 and *S. meliloti* 2011*rhbA62* allowed a clear fragmentation profile for the rhizobactin 1021 molecule to be created. As the level of rhizobactin 1021 was highly concentrated in the *S. meliloti* 2011 extract the rhizobactin 1021 associated ions could be clearly identified in the whole sample through to the MS^3 spectrum. The *S. meliloti* 2011*rhbA62* negative control provided a reference to which the validity of the rhizobactin 1021 spectra could be compared and it was found that the coincidental presence of ions with molecular weights attributed to rhizobactin 1021 in the negative sample was very low, with only three matches present. This gave confidence in the fragmentation pattern for rhizobactin 1021 which allowed the extracts from the *S. meliloti* 2011*rhbD* (pCC103) and *S. meliloti* 2011*sma2339* strains to be examined for rhizobactin 1021. It was clear that the siderophore produced by the *S. meliloti* 2011*sma2339* strain was rhizobactin 1021. There were only slight discrepancies between the spectra observed from the *S. meliloti* 2011 extract and that of the *S. meliloti* 2011*sma2339* which could be attributed to the lesser intensity of the siderophore ions in the mutant extract resulting in greater interference from background noise. The problem of background noise interference was compounded in the *S. meliloti* 2011*rhbD* (pCC103) extract as the signals representing ions associated with rhizobactin 1021 were almost occluded by the background. However the spectra obtained from analysis of this mutant did not resemble the negative control and all the spectra taken together indicate that rhizobactin 1021 was present in the sample but due to such low level being present the result is not conclusive.

Originating from the phenotype of the *sma2339* mutant it is proposed that the *sma2339* gene be officially named *rhbG*, **Rh**izobactin **B**iosynthesis **G**. The data presented here clearly shows that it is involved in the biosynthesis of rhizobactin 1021 as its removal resulted in a 10 fold reduced in siderophore levels. A distinct essential role for this protein could not be identified however it can poorly complement an *rhbD* mutant strain.

The models for rhizobactin 1021 biosynthesis must be slightly altered to accommodate these new results. It would appear that both of the acetyltransferases occupy the same role in the pathway in that each is capable of complementing the others activity. Also

rhizobactin 1021 is still produced when only one of the proteins is present which confirms that they fully overlap in function despite the extra region of unknown function present on the Sma2339 protein. The proposed changes would be to amalgamate the functions of the acetyltransferases in the Lynch model and to leave the role of lipid addition open to speculation. There is no evidence for or against suggesting that the lipid is added to rhizobactin 1021 by the acetyltransferases, as the complete siderophore appears to be formed in all cases. This leaves the possibility that the lipid could be added through the action of a separate protein not transcribed close to the biosynthesis operon. This subtle but important change is shown in figure 3.35 below.



1. L-Glutamic Acid
2. L-Aspartic B-semialdehyde
3. 2-Ketoglutaric Acid
4. L-2,4-diaminobutyric Acid
5. 1,3-diaminopropane
6. N^4-hydroxy-1-aminopropane
7. N^4-acetyl-N^4-hydroxy-1-aminopropane
8. Citrate aminopropane intermediate
9. Schizokinen
10. Rhizobactin 1021

Figure 3.35: Modified rhizobactin 1021 biosynthesis pathway

3.11 Conclusions

The data presented here provide an in depth analysis into the role of RhbD and RhbG in the production of rhizobactin 1021. These data have been summarised and discussed in the previous section. There remain a number of interesting questions in regard to acetyltransferases in rhizobactin 1021 production outside of those addressed previously, the most prominent of which is "Why two acetyltransferases?". To address this question the gene arrangement of the rhizobactin 1021 cluster must be revisited.

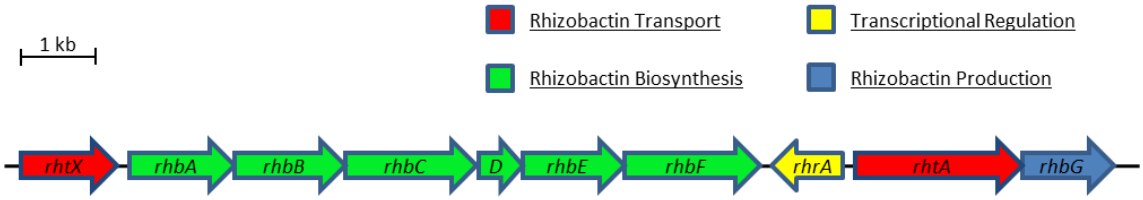


Figure 3.36: Rhizobactin 1021 biosynthesis operon

The position of the two acetyltransferases relative to one another shows that they are transcribed as part of two separate polycistronic transcripts. RhbD is a member of the *rhbABCDEF* polycistron and *rhbG* is a member of the *rhtArhbG* polycistron. This gene arrangement is unusual as both genes fulfil the same function but are transcriptionally separate. The regulation of these genes is also of interest as the transcriptional regulator RhrA has been shown by Lynch et al. (2001) to be the activator of both transcripts. They reported that siderophore production and acquisition was abolished in an *rhrA* transposon mutant. Further to this, unpublished results obtained by Viguier (2005) in which the expression levels of both polycistrons were monitored in an *rhrA* knockout strain in comparison to a wild type strain, found that expression was reduced to 1-3 % of wild type in the mutant. This gene arrangement and strong activation by RhrA would result in an abundance of RNA transcripts for acetyltransferase functions relative to other rhizobactin biosynthesis genes. The reasons for this can only be speculated on but the acetyltransferase function may be a rate limiting step in the production of rhizobactin 1021 and this saturation by dual production of acetyltransferases may overcome the limitation. Indeed this hypothesis would fit with the observed results as production is reduced when rhizobactin production is reliant on one acetyltransferase alone.

A second reason for the transcriptional separation of the genes would be to allow them to be differentially expressed under varying conditions. A microarray study carried out

by Ampe et al. (2003) found that *rhbG*, referred to as *sma2339* was up regulated by the plant flavonoid luteolin. Luteolin is the alfalfa flavonoid responsible for the initiation of symbiosis between *Medicago sativa* and *S. meliloti* (Perret, Staehelin and Broughton 2000). However a second microarray study carried out by Barnett et al. (2004) assessing the effect of luteolin on the whole transcriptome of *S. meliloti* found no such up regulation of *rhbG*. This contradiction may be explained by the variations in the conditions used by each author. Ampe et al. (2003) used a 10 μM final concentration of luteolin and allowed growth to proceed for 12 hr post induction whereas Barnett et al. (2004) induced expression with 3 μM luteolin and allowed 3 hr of growth post induction. These are large variations in culture conditions that would result in the cultures being in different stages of growth and also under different levels of stress from the inducing molecule luteolin. Because of these inconsistencies it is impossible to determine which result is indicative of the true situation found outside of lab conditions, but if *rhbG* is induced by luteolin it could result in an accumulation of rhizobactin 1021 proximal to the plant root in the wild. It is stated previously that rhizobactin 1021 is not essential for symbiosis Lynch et al. (2001) but an increase in siderophore levels proximal to sites of infection would help insure that infecting bacteria would have their iron requirement fulfilled prior to infection. Also high levels of siderophore proximal to a plant wound site have been previously shown to counter the plant's innate immune response. Dellagi et al. (1998) showed that the plant pathogen and causative agent of fire blight, *Erwinia amylovora* produces the siderophore ferrioxamine under iron stress. Studies of the infection rate of a biosynthesis mutant showed that in the absence of the siderophore, infection rates were reduced. It was found that the siderophore was chelating excess iron at the wound site which resulted in the prevention of free radical formation resulting from hydrogen peroxide production by the plant. The siderophore was blocking the reduction of Fe^{3+} to Fe^{2+} by hydrogen peroxide therefore preventing the Fenton reaction from taking place. This would increase the bacterial cell count at the wound site and increase the chances of an infection.

The role of RhbG in the biosynthesis of rhizobactin 1021 outside that of lipid addition was not investigated as part of this study. However unpublished results obtained by Graham Dodrill suggest that RhbG may occupy the role of a scaffold around which the other biosynthesis proteins are localised. The interactions of each of the biosynthesis proteins RhbABCDEFG was mapped against the rest of the biosynthesis proteins in a

pairwise manner by two hybrid analysis. The results obtained by this study demonstrated that RhbG interacts with RhbCDEF, involved in the assembly of rhizobactin 1021 whereas none of the other biosynthesis proteins displayed interaction with more than two. The level of interaction observed by RhbG raises possibilities that it may be acting to localise the biosynthesis proteins in the cell thereby increasing production efficiency. However a more detailed analysis must be undertaken to prove such a role.

The second unresolved issue in relation to rhizobactin 1021 is "what role is the lipid playing?". There are many hypotheses pertaining to the advantages of fatty acid addition to siderophores which span from basic prevention of siderophore pirating to siderophores acting as messenger molecules in biofilm communities.

- From a structural viewpoint the addition of a lipid structure alters the properties of the siderophore along with increasing the size of the siderophore. The larger structure may sterically hinder the siderophore from travelling through a transporter designed just to transport the core non-lipid form. This would prevent the siderophore from being acquired by a neighbouring strain producing a similar siderophore. Also the lipidated siderophore moves from being hydrophilic to being amphipathic resulting from the addition of the hydrophobic carbon chain. This can have important consequences for transport of the siderophore. TonB-dependant receptors often identify their targets through just 3-4 residues and any structural change in the substrate could disrupt this interaction or indeed transport by inner membrane transporters (Koebnik, Locher and Van Gelder 2000). Such a disruption was observed by O Cuiv et al. (2007) when investigating ferrichrome, ferrioxamine B, schizokinen and rhizobactin 1021 transport across the inner membrane by the protein FoxB. It was found through complementation that FoxB was capable of transporting schizokinen but could not transport rhizobactin 1021. This failure to transport can be attributed solely to the presence of the lipid as it is the only difference between the two siderophores.
- Lipid structures on siderophores have been shown to increase the siderophore's affinity for cellular membranes. Studies carried out on acinetoferrin in which the membrane affinity of the siderophore to vesicles were assessed indicated a strong membrane affinity for the iron free siderophore and a weaker affinity for

the membrane upon iron binding (Luo, Fadeev and Groves 2005). This would allow the producing organism to localise the siderophore to its surface and upon binding iron result in the release of the siderophore allowing transport to occur. It would also help prevent the siderophore from being scavenged by a nearby organism as it would help prevent diffusion. Prevention of diffusion would be of paramount importance in an aqueous environment as it would occur more rapidly than in a complex soil environment. This is reflected by the variety of lipidated siderophores from marine bacteria with the marinobactins, synechobactins and ochrobactins being well studied examples. Membrane association studies on these siderophores have shown that each siderophore can attach itself to vesicles and in the case of ochrobactins to the cell surface of the producing organism (Martin et al. 2006; Martinez et al. 2003). In addition numerous lipidated siderophores from marine bacteria possess the ability to assemble into micelle formations (Martinez et al. 2000).

- Two comparative metabolomic studies carried out in *Streptomyces coelicolor* found novel conditions under which an extensive suite of deferentially lipidated ferrioxamines were produced. The first study assessed the changes in secondary metabolites between a culture grown at the standard temperature of 28 °C and a stress temperature of 37 °C. It was found that this temperature stress induced the production of 15 amphipathic ferrioxamines, a number of which had never been observed previously (Sidebottom et al. 2013). The acyl chain length on these siderophores varied in length between 11 carbons and 15 carbons. The second study in *S. coelicolor* assessed metabolomic changes when grown in proximity to other strains of actinomycetes. It was found that at least 12 different acyl-ferrioxamines were produced and it was found that iron competition between the strains was the initiator of production (Traxler et al. 2013). It was observed that when grown in the proximity of a siderophore producing strain, *Amycolatopsis sp.* AA4, the full array of siderophores were produced but in a siderophore deficient mutant of the same strain only four was produced. In this case the production of such a wide variety of siderophore structures appears to be a defence against siderophore pirating as the presence of a siderophore producing strain resulted in an increased variety. The production of siderophores in response to heat stress is less straight forward. It was found that very low quantities were produced 170 ng/l in comparison to ~20 mg/l from *S. meliloti*

2011, which suggests that these siderophores are not being produced to counteract an iron stress. Perhaps it is a strategy to inhibit the growth of competing strains in times of stress as this would increase the chance of survival of the producing strain.

As discussed there are a wide variety of benefits to producing a lipidated siderophore. It is unlikely that *S. meliloti* 2011 would have to contend with environmental conditions similar to the aqueous environments encountered by the marine bacteria that results it prohibitive rates of diffusion. Perhaps rhizobactin 1021 is lipidated as a means to counter its utilisation by surrounding strains therefore allowing the producing strain to stockpile a siderophore reserve for its sole use.

Chapter Four

Coprogen Mediated Iron Acquisition in *Sinorhizobium meliloti* 2011

4.1 Introduction

The ability of an organism to partake in siderophore pirating represents a relatively low energy technique for acquiring iron and also provides a competitive edge over neighbouring organisms. It allows for the acquisition of sufficient levels of iron without the metabolic cost of producing an endogenous siderophore. *S. meliloti* 2011 has been shown to utilise a variety of xenosiderophores namely schizokinen, ferrichrome and desferrioxamine B (O Cuiv et al. 2008; O Cuiv et al. 2004). Uptake of schizokinen was found to occur through the rhizobactin 1021 transport system. The TonB-dependent outer membrane receptor RhtA allows for the internalisation of the siderophore into the periplasmic space where it is then transported across the inner membrane by the major facilitator superfamily transporter RhtX (O Cuiv et al. 2004; Lynch et al. 2001). Figure 4.1 is a representation of schizokinen and rhizobactin 1021 uptake.

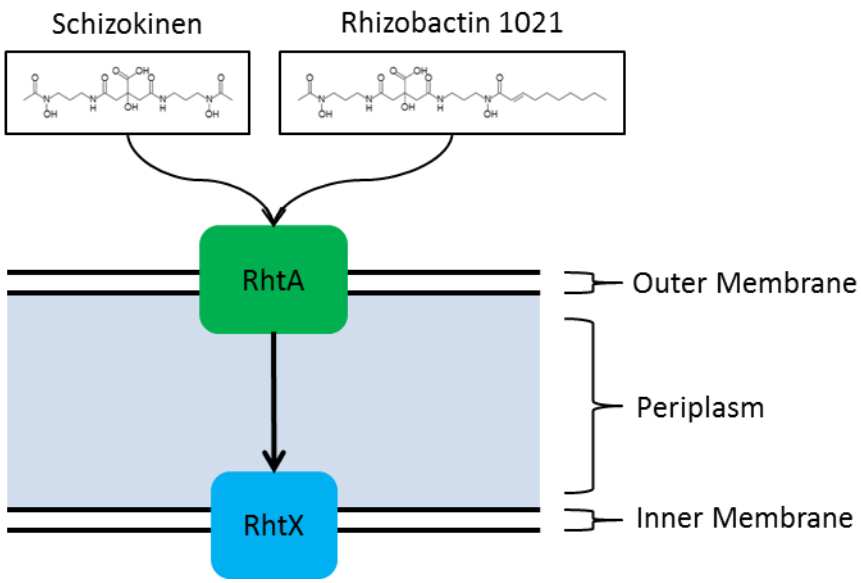


Figure 4.1: Schematic of rhizobactin 1021 and schizokinen uptake

Ferrichrome and ferrioxamine B share a number of transport proteins with each other and with that of haem acquisition in a unique shared "split system" (O Cuiv et al. 2008). Each siderophore has its own cognate TonB-dependent outer membrane receptor, FhuA and FoxA for ferrichrome and ferrioxamine B respectively. Once internalised to the periplasm the siderophores are bound to the periplasmic binding protein FhuP. The siderophores are then transported to the inner membrane ABC transporter complex for internalisation. The inner membrane complex comprises the HmuU permease and HmuV ATPase which are generally associated with haem uptake.

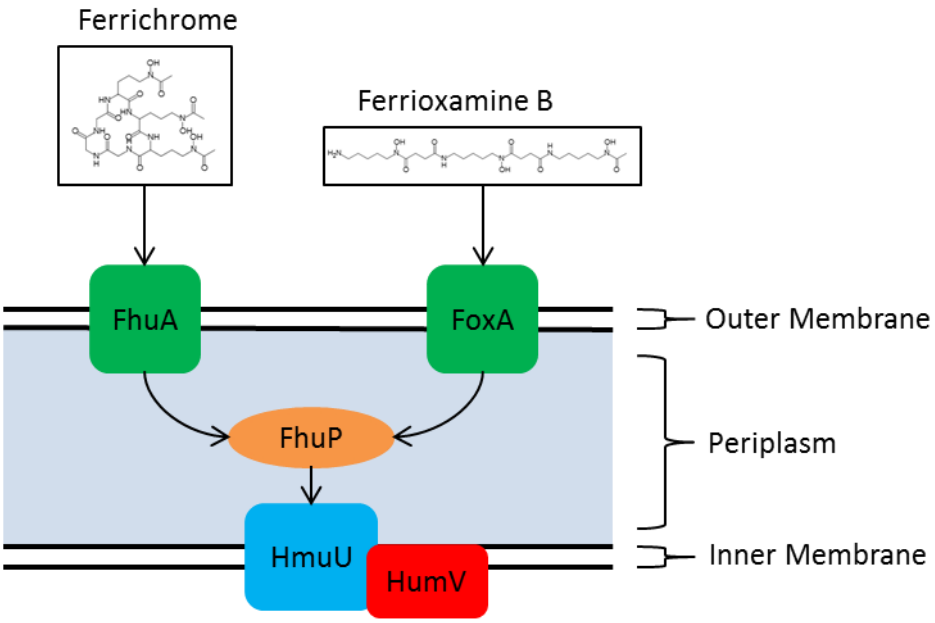


Figure 4.2: Schematic of ferrichrome and ferrioxamine B uptake

The genome structure of *S. meliloti* consists of three independent replicons; the chromosome with a length of 3.65 Mb and two megaplasmids pSymA and pSymB with respective lengths of 1.35 Mb and 1.68 Mb. It is worth mention that the outer membrane receptors FhuA, FoxA and ShmR required for the uptake of ferrichrome, ferrioxamine B and haem respectively are transcribed from genes located on the chromosome with *rhtA* located on the pSymA megaplasmid. In addition to *rhtA* there is one other putative outer membrane receptor located on pSymA which is designated *sma1747*. This arrangement is represented in figure 4.3.

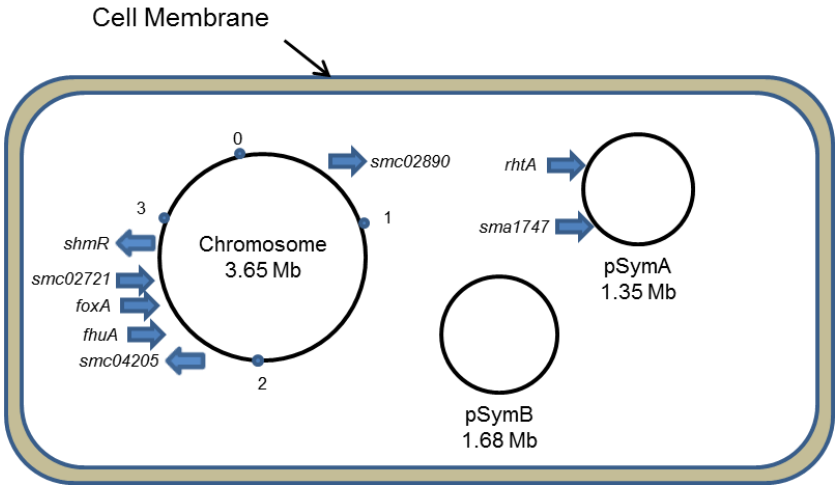


Figure 4.3: Representation of the genome structure of *S. meliloti* 2011

Unpublished work carried out by Dr. Damien Keogh assessed the breadth of xenosiderophore acquisition in *S. meliloti* 2011. In addition to testing ferrichrome and ferrioxamine B the ability to utilise coprogen was assessed. This screen was carried out using the iron nutrition bioassay described in Lynch et al. (2001) which allowed many siderophores to be tested relatively quickly with little experimental set up. As part of this analysis a number of *S. meliloti* 2011 mutants and derivatives were assessed namely *S. meliloti* 2011, *S. meliloti* 2011*rhtX-3, S. meliloti* 2011*rhbA62,* and *S. meliloti* Rm818 and the genotype of each strain is documented in Chapter 2. In brief, this analysis showed that *S. meliloti* 2011 is capable of utilising coprogen as an iron source. Also observed was that the strain *S. meliloti* Rm818 was deficient for utilisation. This allowed for at least some elements of the coprogen utilisation pathway to be mapped to the pSymA megaplasmid. The only characterised uptake system on pSymA is the rhizobactin 1021 uptake pathway. This was ruled out as the coprogen uptake pathway as both the strains *S. meliloti* 2011*rhtX-3* and *S. meliloti* 2011*rhbA62* were positive for utilisation.

The screen of coprogen utilisation described above maps essential genes involved in coprogen transport to the pSymA megaplasmid. The results contained in this chapter describe the assessment of the likely candidates for the coprogen utilisation system.

4.2 Bioinformatic analysis of the pSymA megaplasmid for candidate coprogen receptors

Numerous methods of siderophore transport across the inner membrane have been described as outlined by Schalk and Guillon (2013). However the only method described for transporting siderophores through the outer membrane is via TonB-dependent outer membrane receptors. Based on this uniform transport mechanism it was decided to assess the *S. meliloti* genome for outer membrane receptors located on the pSymA megaplasmid using the BLASTp program.

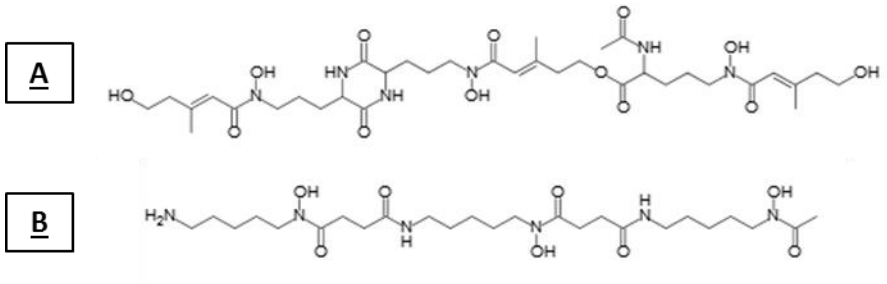


Figure 4.4: Structure of coprogen and ferrioxamine B, labelled A and B respectively

It was decided that the sequence of the FoxA protein from *S. meliloti* 2011 would be the most prudent subject sequence with which to analyse pSymA for a possible coprogen receptor. Ferrioxamine B and coprogen are both hydroxymate siderophores and have been found to be transported through the FhuDBC system in *E. coli* (Clarke et al. 2002). Resulting from this overlap in transport it would be expected that their outer membrane transporters would show a high level of similarity. Equally the outer membrane receptor for ferrichrome, FhuA could have been used in this section as ferrichrome is also transported through the FhuDBC inner membrane system in *E. coli*.

Table 4.1: *S. meliloti* 2011 proteins similar to FoxA as determined by BLASTp

Protein	Genome Location	Query Coverage (%)	Identity (%)	E Value
FhuA	Chromosome	99	33	5e^-111
Sma1747	pSymA	99	31	7e^-88
RhtA	pSymA	29	23	0.003
ShmR	Chromosome	41	23	0.004

As expected using FoxA as the subject sequence the BLASTp search identified FhuA as the most significant homologue. However, FhuA is located on the chromosome and it cannot be the receptor for coprogen. Sma1747 shows nearly the same level of identity to

FoxA as FhuA and is located on pSymA. Resulting from this Sma1747 was selected as the most probable outer membrane receptor for coprogen. To confirm that the results shown in table 4.1 are not a result of each protein simply being a TonB-dependent outer membrane receptor the protein sequences were aligned using the Multiple Sequence Comparison by Log-Expectation (MUSCLE) algorithm. Figure 4.5 and figure 4.6 show alignments of the FoxA, FhuA, and Sma1747 proteins and Sma1747, ShmR, and RhtA proteins respectively. Alignments are coloured based on percentage identity at each locus with dark blue representing 100% conservation, light blue 66% conservation and no colour no conservation. Sequence gaps are represented by a hyphen.

As can be seen in figure 4.5 it is clear that Sma1747 is very similar to FoxA and FhuA across the entire length of the protein as shown by the level of 100% conservation and by few gaps being needed to force the alignment. In correlation with this figure 4.6 shows the Sma1747 is dissimilar to ShmR and RhtA over the full length of the protein with the alignments interrupted by a large number of gaps to make this alignment fit.

This analysis taken along with the BLASTp results, indicates that Sma1747 is closely related to the hydroxymate TonB-dependent outer membrane receptors which further strengthens the probability of it being the receptor for coprogen.

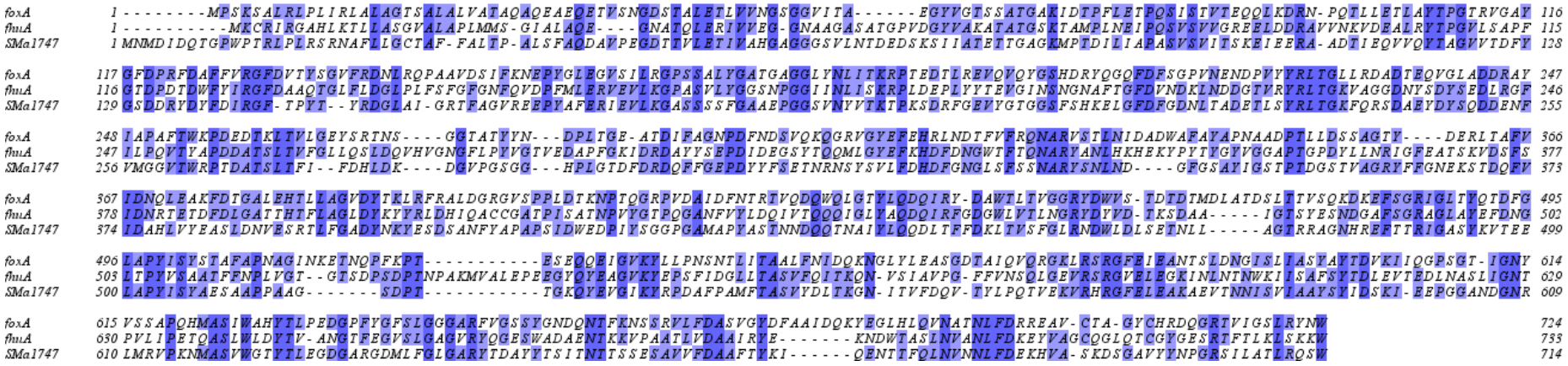


Figure 4.5: Protein sequence alignment of FoxA, FhuA and Sma1747

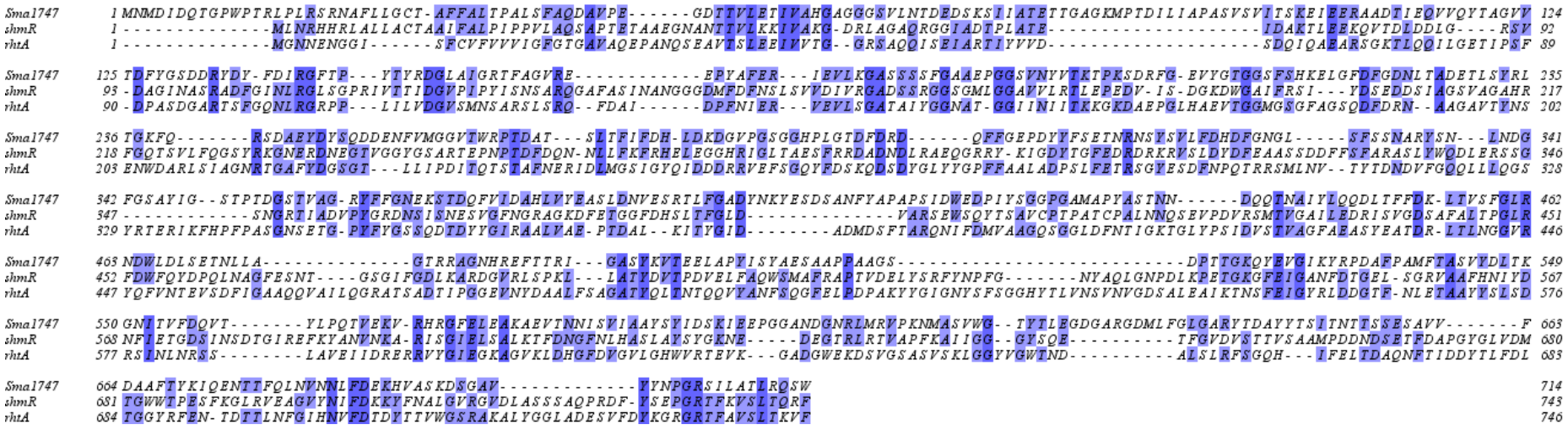


Figure 4.6: Protein sequence alignment of Sma1747, ShmR and RhtA

4.2.1 Analysis of the genome proximal to *sma1747*

Observation of the genome region downstream to *sma1747* reveals a possible operon of genes, *sma1746-sma1740*, annotated as a Fe^{3+} transport system with the terminal gene annotated as a siderophore interacting protein. The transport system has a typical ABC transporter arrangement comprising a periplasmic binding protein, two inner membrane permeases and an ATPase protein. Upstream of *sma1747* there is a divergently encoded transcriptional regulator *sma1749* and a large intergenic region followed by a hypothetical protein of unknown function.

Bioinformatic analysis of Sma1749

The gene *sma1749* is located at position 991460-992359 of the pSymA megaplasmid. The encoded protein is predicted to be 299 amino acids in length with a molecular weight of 33 kDa and a pI of 9.35. The amino acid sequence of the predicted Sma1749 protein is given in figure 4.7 below.

```
MGFWHSMSWKTEGIRVTAPVQWRQYDGMVSVLWEAESQAGASGYYLADDP
RIMFFFNDVSSSVRISNQDNDLARNSRPMARAVYVPAGVPLWTGSRKTHS
FSHLNLHLHKDRVLRFLAPSVGNSAAQTALRCPVELQNVAAIDALAKLLA
DEIKSPTKHAVYAESLIGSIIAGLLDIPTEREERAGGRLTQAQINKLVSH
IDALGDYRMSVADMAAVVGLSESWFASVFKQTMGITPLQWQLAKRVELTK
KLLGESDLPVASIAAQLGFADQAHLTKVFRQIVGETPAAWRRMRQFRQS
```

Figure 4.7: Amino acid sequence of Sma1749

This protein sequence was analysed by the InterPro web tool to allow the protein to be classified based on family, conserved domains and conserved sites (Hunter et al. 2012). A distinct helix turn helix domain was identified at the C-terminus of the protein similar to that found in the AraC-type DNA binding proteins. The N-terminus shows no identifiable domains or sites but this is usual for transcriptional regulators as the N-terminus is associated with protein dimerization and substrate binding.

Bioinformatic analysis of sma1747

The gene *sma1747* is located at position 989276-991420 of the pSymA megaplasmid. The encoded protein is predicted to be 714 amino acids in length with a molecular weight of 78 kDa and a pI of 4.51. The amino acid sequence of the predicted Sma1747 protein is given in figure 4.8 below.

MNMDIDQTGPWPTRLPLRSRNAFLLGCTAFFALTPALSFAQDAVPEGDTT
VLETIVAHGAGGGSVLNTDEDSKSIIATETTGAGKMPTDILIAPASVSVI
TSKEIEERAADTIEQVVQYTAGVVTDFYGSDDRYDYFDIRGFTPYTYRDG
LAIGRTFAGVREEPYAFERIEVLKGASSSSFGAAEPGGSVNYVTKTPKSD
RFGEVYGTGGSFSHKELGFDFGDNLTADETLSYRLTGKFQRSDAEYDYSQ
DDENFVMGGVTWRPTDATSLTFIFDHLDKDGVPGSGGHPLGTDFDRDQFF
GEPDYYFSETNRNSYSVLFDHDFGNGLSFSSNARYSNLNDGFGSAYIGST
PTDGSTVAGRYFFGNEKSTDQFVIDAHLVYEASLDNVESRTLFGADYNKY
ESDSANFYAPAPSIDWEDPIYSGGPGAMAPYASTNNDQQTNAIYLQQDLT
FFDKLTVSFGLRNDWLDLSETNLLAGTRRAGNHREFTTRIGASYKVTEEL
APYISYAESAAPPAAGSDPTTGKQYEVGIKYRPDAFPAMFTASVYDLTKG
NITVFDQVTYLPQTVEKVRHRGFELEAKAEVTNNISVIAAYSYIDSKIEE
PGGANDGNRLMRVPKNMASVWGTYTLEGDGARGDMLFGLGARYTDAYYTS
ITNTTSSESAVVFDAAFTYKIQENTTFQLNVNNLFDEKHVASKDSGAVYY
NPGRSILATLRQSW

Figure 4.8: Amino Acid sequence of Sma1747

Analysis of Sma1747 by the InterPro web tool revealed that it is a member of the TonB-dependent siderophore receptor family. The two main domains identified were the N-terminal plug region at position 71-195 and the beta-barrel region at position 203-714. These two regions are found in all known TonB-dependent siderophore receptors. In addition to these two domains a signal peptide sequence was identified at position 1-40 indicating that the protein is exported out of the cytoplasm.

Bioinformatic analysis of Sma1746

The gene *sma1746* is located at position 987982-988983 on the pSymA megaplasmid. The encoded protein is predicted to be 333 amino acids in length with a molecular weight of 36 kDa and a pI of 4.78. The amino acid sequence of the predicted Sma1746 protein is given in figure 4.9.

MAMRLFRFAVLAVIALHWADRHGTVFAAESTSYPITIKHAFGNTIIAKKP
ERVATVAWANHEVPLALGIVPVGMARANFGDDDDDGILPWVDARLGELKA
EKPMLFDEGDGIDFEAVAATRPDVILAAYSGLSQADYDTLSDIAPVIAYP
QAPWSTDWRETIRLNSAGLGMAAEGEGLIASIEAEIDLALDGHPELKGKS
AMFITHLSSWDLSVVNFYTTNDTRVRFFGDLGLMSPKSVVQASQPGRFSG
SVSAEQIDAFDDVDILVTYGDGMLFDALKANALMMHMPAVARESIVMLGN
NAVGNAANPTPLSIRWVLKDYVKLLSEAAKKSQ

Figure 4.9: Amino acid sequence of Sma1746

Analysis of Sma1746 by the InterPro web tool showed no definite membership of a protein family. The domain analysis found an ABC transporter periplasmic binding domain covering nearly the full length of the protein at position 52-329 with a signature match to iron and vitamin B12 periplasmic binding proteins. A signal peptide sequence was also identified at position 1-27 indicating that the protein is exported to the periplasm.

Bioinformatic analysis of Sma1745

The gene *sma1745* is located at position 986939-988039 on the pSymA megaplasmid. The encoded protein is predicted to be 366 amino acids in length with a molecular weight of 37 kDa and a pI of 9.95. The amino acid sequence of the predicted Sma1745 protein is given in figure 4.10.

```
MGPEGLREAAQRGSQEVAMTAGIIKTSDGRWSPVSRSNRMRSLWLAALFA
LVIVLLAASVTVGTRNVGWDDVAAAMGGAQNNIDRASVALRIPRTVLALL
AGGALGLAGAIMQGVTRNPLADPGILGVNMGASLAVVVGVAWFGMSSLYA
YIWVAILGAGCAAIFVYAIGSLGRGGATPLKMALAGAATSVAFASMVIAV
VLPRGDIAGGIHSWQIGGVGGATYERILPVLPFLLVGFVISLLSARKLNS
LALGDELATGLGESVATARAVASLGAILLCGATTAICGPIGFLGLVVPHL
CRLLVGVDHRWLLPFSTLAGACLLLAADVLGRIVARPAEIDVGIVTAMIG
APFFIWIVRRQRVRDL
```

Figure 4.10: Amino acid sequence of Sma1745

Analysis of Sma1745 by the InterPro web tool classified it as a member of the ABC transporter permease protein family. No distinct domains were identified however a signature match was identified to FecCD, the permease unit of the ABC transporter for dicitrate uptake in *E. coli*.

Bioinformatic analysis of Sma1742

The gene *sma1742* is located at position 985905-986942 on the pSymA megaplasmid. The encoded protein is predicted to be 345 amino acids in length with a molecular weight of 35 kDa and a pI of 11.38. The amino acid sequence of the predicted Sma1742 protein is given in figure 4.11.

```
MTAPSSTLAVVIAHRRKRARRHHAIIATLLTLVAVTFGVTLSIGQSITPP
SDVLRVLLGEPVPGASFTVGQLRLPRAVLSILAGLCFGLGGVAFQVMLRN
PLASPDIIGITSGAGAAAVFAIVVLSMTGPMVSVIAVVAGLGVALLVYAL
SFRNGVAGTRLILVGIGVSAMLQSVIAYILQSAPAWNLQEAMRWLTGSVN
GAQLGQALPLLLALIFFGGLLLVRGRDLETLRLGDDTAAALGTRVSNTRM
LVIVAAVGLIASATAASGPIAFVAFLSGPIAGRIVRNDGSVLIPSALTGA
VLVLAADYVGQHLLPSRYPVGVVTGALGAPYLLYLIVRINRIGGS
```

Figure 4.11: Amino acid sequence of Sma1742

The InterPro analysis of Sma1742 yielded the same results as for Sma1745. The protein is a member of the ABC transporter system, permease component and displays a signature sequence match to FecCD from *E. coli*.

Bioinformatic analysis of Sma1741

The gene *sma1741* is located at position 985063-985905 on the pSymA megsplasmid. The encoded protein is predicted to be 280 amino acids in length with a molecular weight of 30 kDa and a pI of 6.25. The amino acid sequence of the predicted Sma1741 protein is given in figure 4.12.

```
MTDHLLVASGLTAGYDKTEILHALDLTIPPRKITVIVGANACGKSTFLRT
LSRLIAPSKGQVLLDGKSIHRTPSRDLARTLGLLPQSPIAPEGITVVDLV
SRGRHPHQSLFSGWTRRDDEAVDSALRATKTFDLADRPIDELSGGQRQRV
WIAMALAQQTDILLLDEPTTFLDINHQIEVLDLLTDLNSARRTTVVMVLH
DLNLAARYADHLVAIAGGRVHISGTPEEVLTEETVRHVFGLDSRVISDPT
SGRPIMLPIGRHRTAVIDDMGDAPQKERSA
```

Figure 4.12: Amino acid sequence of Sma1741

The InterPro analysis of Sma1741 found no protein family membership. However a number of signature matches were observed: a P-loop containing nucleotide triphosphate hydrolase match was found at position 4-234, a AAA+ ATPase domain match was identified at position 30-219 and a ATP-binding cassette with a ABC transporter-type domain profile was observed at position 6-242. These matches strongly indicate that Sma1741 is an ATPase involved in inner membrane transport.

Bioinformatic analysis of Sma1740

The gene *sma1740* is located at position 983996-985066 of the pSymA megaplasmid. The encoded protein is predicted to be 356 amino acids in length with a molecular weight of 39 kDa and a pI of 5.11. The amino acid sequence of the predicted Sma1740 protein is given in figure 4.13.

```
MNAVHSFKLSGVAVPGSAADMLDEICEHFVEHAKVERRDDLAVLQSELGV
ARISIENGRLLIELDCPTREKLHMSRTILAEHLFYFAEGQPFELTWSEPT
SLSVLPNLHEVTVVSAHDVTPHMRRVIFSCVDVTPFVGSDMHVRLLVPPK
GKPPVWPGYREDGRIAWPEGENELLVRVYTIRAVDLDRSELCIDFLQHPA
PGVPTPGADFARDAQPGDVAALLGPGAGGLPAERSILLIGDESALPAIAR
IAAEAPAETHIRAIIEVEDKAEEQPLLTDGVLDVRWLHRGSYPGDAADIL
VSEAKAAISAVDDETFVWVACERTDIRAIRTFLKARQHDRRKMYVAWYWE
RDVKIA
```

Figure 4.13: Amino acid sequence of Sma1740

The InterPro analysis of Sma1740 classified it as a member of an uncharacterised conserved family of proteins. Two distinct domains were identified with the first at position 99-232 annotated as a ferredoxin reductase-type FAD-binding domain and the second a siderophore-interacting domain at position 234-351. The identified signature

matches all correlate with the two identified domains. These results suggest that Sma1740 is a siderophore reductase responsible for the removal of iron from a siderophore-Fe^{3+} complex by reduction of Fe^{3+} to Fe^{2+}.

4.2.2 Summary of the analysis of the region surrounding *sma1747*

The analysis of the proteins encoded proximal to *sma1747* shows that there appears to be a complete transport system for siderophore uptake encoded at that genome location. Based on bioinformatic analysis the genes *sma1746-sma1741* encode for an ABC transporter complex comprising a periplasmic binding protein, two inner membrane permeases and an associated ATPase. Analysis of Sma1740 strongly suggests that it is a reductase protein possibly involved in the release of iron by reduction of ferric-siderophore complexes. The positioning of the Ara-C type regulator *sma1749* adjacent to *sma1747* suggests that it is responsible for initiating transcription of *sma1747* under the correct conditions, possibly in the presence of the cognate siderophore. Figure 4.14 is a diagrammatic representation of the possible coprogen uptake operon. Note that there are no genes annotated as *sma1743*, *sma1744* or *sma1748* in the *S. meliloti* genome.

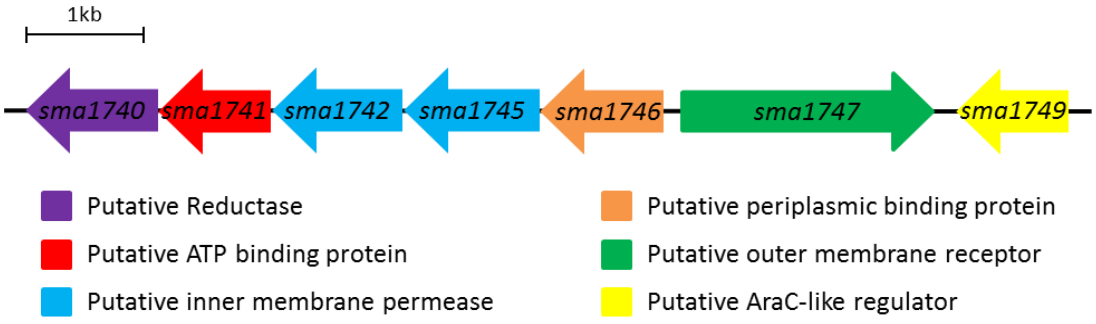


Figure 4.14: Gene arrangement of the possible coprogen uptake system

Further to this analysis, results obtained by Chao et al. (2005) while assessing the role of the global transcriptional regulator RirA confirm that this region is iron regulated. They found that under iron limitation the genes *sma1746*, *sma1745* and *sma1749* are induced. It was found that *sma1747* was not induced under iron limiting conditions which may indicate a separate mechanism of regulation other than iron responsiveness, perhaps mediated by the AraC-like protein Sma1749. Perhaps Sma1749 also regulates *sma1742*, *sma1741* and *sma1740* as they do not appear to be induced under iron stress.

However this is unlikely as there are very short intergenic regions between each of these genes.

The focus of the remainder of this chapter will be to assess Sma1747 for transport of the siderophore coprogen through complementation of *S. meliloti* Rm818 and the creation of a knockout mutant by allelic exchange. Having confirmed Sma1747 as the outer membrane receptor the mechanism of transport across the inner membrane will be assessed. It was decided to carry out this analysis without consideration of the *sma1740* gene. This was done as previous work carried out on possible reductases demonstrated that they are either not required by *S. meliloti* 2011 or there is redundancy present due to related functions from other systems (O Cuiv et al. 2008).

4.3 Complementation of *S. meliloti* Rm818 with *sma1747* and *sma1746-sma1741*

As discussed previously *S. meliloti* Rm818 is unable to utilise coprogen as an iron source due to being cured of the pSymA megaplasmid. This provides an excellent background to carry out complementation studies using the *sma1747, sma1746-sma1741* transport system. As the pSymA plasmid is abolished coprogen utilisation in the complemented strain will show that this is the full transport system for coprogen.

4.3.1 Cloning of *sma1747*

The *sma1747* gene including the ribosome binding site was cloned into the broad host range vector pRK415. Two primers were designed to amplify *sma1747* from genomic DNA prepared from *S. meliloti* 2011. The forward primer was designed to include the ribosome binding site upstream of the start codon and to incorporate an EcoRI site at the 5' end of the PCR product. The reverse primer was designed to include the stop codon of *sma1747* and to incorporate an EcoRI site at the 3' end of the PCR product. The sequences of each primer are given in table 4.2. Using the conditions detailed in table 4.3 below *sma1747* was amplified from *S. meliloti* 2011 genomic DNA. The PCR product was purified by gel extraction to remove any non-specific PCR products and the resulting purified DNA was digested using the EcoRI restriction enzyme. A previously prepared pRK415 vector was also digested using the EcoRI restriction enzyme. Each DNA product was then purified and used in a ligation reaction. Following transformation a plasmid screen was carried out until an approximate 2.2 kb increase was observed in the vector. This was subsequently analysed by restriction digest to confirm the insert was the *sma1747* amplified gene. A correct clone was isolated and named pKC101.

Table 4.2: Primer sequences for cloning *sma1747* into pRK415

Name	Sequence
Sma1747F	ATGCCATTCGAATTCTACGTCAATTGGAACATGAAC
Sma1747R	ATGCCATTCGAATTCTTACCAACTCTGGCGCAGG

Table 4.3: PCR conditions for amplification of *sma1747*

PCR Cycling Conditions	
Annealing Temperature	60 °C
Annealing Time	15 sec
Extension Time	45 sec

The plasmid pKC101 was introduced into *S. meliloti* Rm818 by triparental mating. This background was then used for the introduction of the *sma1746-sma1741* genes *in trans*.

4.3.2 Cloning of *sma1746-sma1741*

The *sma1746-sma1741* genes were cloned into the broad host range vector pBBR1MCS-5. Two primers were designed to amplify the genes from the genome of *S. meliloti* 2011. The forward primer was designed to include the ribosome binding site of *sma1746* and introduced an ApaI site onto the 5' end of the PCR product. The reverse primer was designed to include the stop codon of *sma1741* and introduced a HindIII site onto the 3' end of the PCR product. The sequences of each primer are given in table 4.4 below. Using the conditions outlined in table 4.5 the *sma1746-sma1741* genes were amplified from *S. meliloti* 2011 genomic DNA. The PCR product was purified by gel extraction to remove any non-specific products and the purified DNA used in a restriction digest using the ApaI and HindIII restriction enzymes. A previously prepared pBBR1MCS-5 vector was also digested with ApaI and HindIII. The resulting DNA was subsequently purified and used in for a ligation reaction. Following transformation a plasmid screen was carried out until a size shift of approximately 3.9 kb was observed. This was analysed by restriction digestion to confirm that the insert was the *sma1746-sma1741* gene cluster. A correct clone was identified and named pKC102.

Table 4.4: Primer sequences for the cloning of *sma1746-sma1741* into pBBR1MCS-5

Name	Sequence
Sma1746-41F	ATGCCATTCGGGCCCAACGTTGAATGGAGTGTGGC
Sma1746-41R	ATGCCATTCAAGCTTCGCTCCTCTCCTTTTGTGG

Table 4.5: PCR conditions for the amplification of *sma1746-sma1741*

PCR Cycling Conditions	
Annealing Temperature	62 °C
Annealing Time	15 sec
Extension Time	2 min

The plasmid was then introduced into *S. meliloti* Rm818 and *S. meliloti* Rm818 (pKC101) by triparental mating.

4.3.3 Coprogen utilisation by *S. meliloti* Rm818 complemented strains

The following strains were analysed by the iron nutrition bioassay: *S. meliloti* Rm818, *S. meliloti* Rm818 (pKC101), *S. meliloti* Rm818 (pKC102), and *S. meliloti* Rm818 (pKC101, pKC102). *S. meliloti* 2011*rhbA62* was used as a positive control as it gives very clear haloes for iron nutrition bioassays. A single colony from each strain was used to inoculate a 5 ml aliquot of TY broth supplemented with antibiotics where appropriate. Growth was allowed to proceed until early stationary phase, ~2 days at which point 250 µl was used to inoculate 25 ml of molten agar. While the agar was still molten appropriate antibiotics were added along with 2,2'dipyridyl to a final concentration of 300 µM and IPTG to a final concentration of 500 µM. The agar was then poured and allowed to set. Wells were cut in the agar and the following test solutions added: $FeCl_3$ as an inoculum control, ferrioxamine B as a siderophore acquisition control, dH_2O as a negative control and coprogen as the test solution. Coprogen and ferrioxamine B were used at a final concentration of 100 µM. Table 4.6 shows the results of the iron nutrition bioassay experiments.

Table 4.6: Iron nutrition bioassay analysis of complemented *S. meliloti* Rm818

Strain	Plasmid	Ferrioxamine B	Coprogen
***S. meliloti* Rm818**	—	+	—
***S. meliloti* Rm818**	pKC101	+	+
***S. meliloti* Rm818**	pKC102	+	—
***S. meliloti* Rm818**	pKC101 pKC102	+	+
S. meliloti* 2011*rhbA62	—	+	+

Strains were negative for growth promotion by dH_2O.

All strains showed the expected behaviour around each of the control wells. Unexpected results from any of the controls will invalidate the experiment. The result observed in the *S. meliloti* Rm818 strain confirms the previously reported results as it is unable to acquire coprogen. Utilisation of coprogen by *S. meliloti* 2011*rhbA62* indicates that there is sufficient coprogen to stimulate growth.

The dual complemented strain *S. meliloti* Rm818 (pKC101, pKC102) displays the expected phenotype from this analysis as it can acquire coprogen as an iron source. This result suggests that the *sma1747-sma1741* region encodes for proteins capable of transporting coprogen. Also the *S. meliloti* Rm818 (pKC102) phenotype is in keeping with *sma1747* being the outer membrane receptor for coprogen as no growth promotion is observed in the absence of pKC101. However the result observed in the *S. meliloti* Rm818 (pKC101) strain deviates from the hypothesis that *sma1746-sma1741* comprises the sole inner membrane transport system as growth promotion was observed without the presence of the *sma1746-sma1741* genes. This result negates the result observed in the dual complemented strain as the *sma1746-sma1741* is not essential for the acquisition of coprogen. The consequences of this are discussed below.

The results shown in table 4.6 above demonstrate that *sma1747* is the TonB-dependent outer membrane receptor for coprogen. However to confirm that it is the sole outer membrane protein capable of transporting coprogen a knockout mutant of *sma1747* will be made in the *S. meliloti* 2011*rhtX-3* strain. This strain was selected over *S. meliloti* 2011 as it is deficient for rhizobactin 1021 transport and production and therefore provides an excellent background in which to conduct iron nutrition bioassays.

The result observed in the *S. meliloti* Rm818 (pKC101) strain indicates one of two possibilities. The first is that *sma1746-sma1741* does not encode for the inner membrane transport complex for coprogen and the actual inner membrane system is transcribed elsewhere on the genome. Or secondly it does encode for a coprogen transport system and there is a redundant system located elsewhere on the genome. To assess whether *sma1746-sma1741* encodes for transport of coprogen a knockout mutant of *sma1746* will be made in the *S. meliloti* 2011*rhtX-3* background. *S. meliloti* 2011*rhtX-3* was selected for the same reasons as outlined above. Identification of the redundant system may not be possible as there are numerous inner membrane transport systems in *S. meliloti* 2011. However based on previous work carried out by O Cuiv et

al. (2008) in which the protein FhuP of *S. meliloti* 2011 was found to bind both ferrichrome and ferrioxamine B along with evidence from Clarke et al. (2002) that ferrichrome, ferrioxamine B and coprogen can be bound by the *E. coli* protein FhuD it would be logical to assign FhuP as the most likely candidate for the redundant periplasmic binding protein. With this in mind a *fhuP* knockout will be made in a *S. meliloti* Rm818 (pKC101) background. *S. meliloti* Rm818 (pKC101) was selected as it would simulate a double mutant background with the same phenotype as a *S. meliloti* 2011*rhtX-3sma1746* mutant in respect of coprogen utilisation. This is desirable as the other option will be to create a *S. meliloti* 2011*rhtX-3fhuPsma1746* triple gene knockout mutant which may be very sickly due to the presence of three antibiotic resistance cassettes and would also be very time consuming to generate.

4.4 Antibiotic resistance cassette mutagenesis of *sma1747*

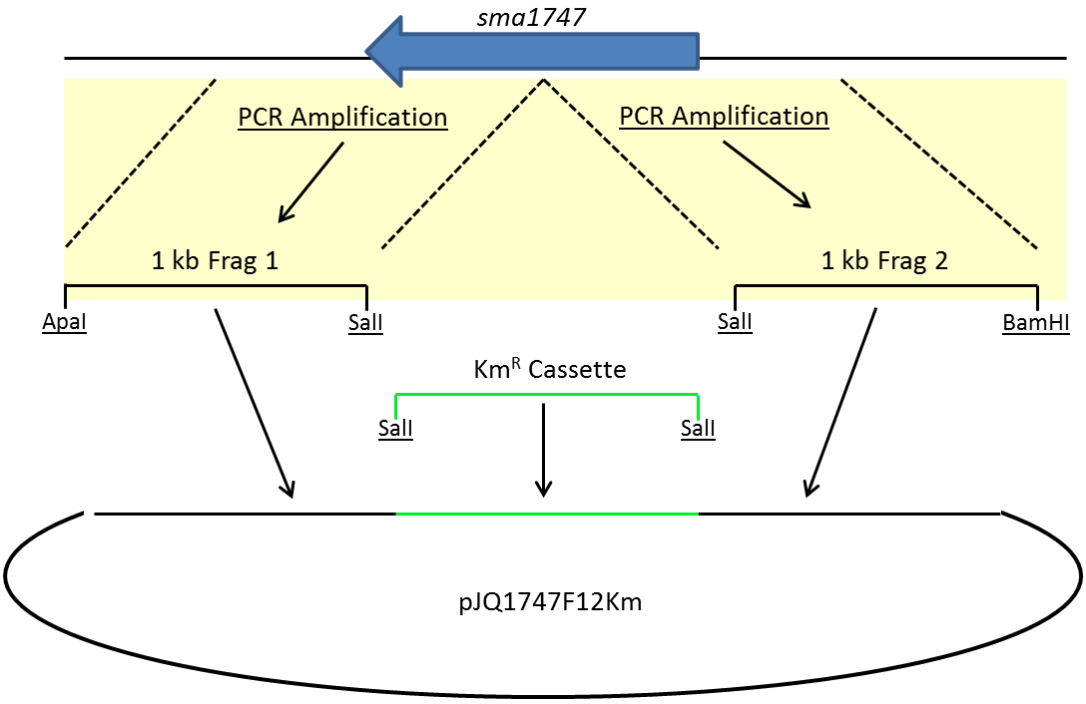


Figure 4.15: Schematic for the creation of the construct used to mutagenize *sma1747*

A gene knockout of *sma1747* was made by disruption with a kanamycin resistance cassette. A brief overview of the process used to create the pJQ200sk based construct to mutagenize *sma1747* is shown in figure 4.15. The region internal to and surrounding *sma1747* was assessed for restriction sites amenable for the introduction of a kanamycin resistance cassette and no naturally occurring sites were present. Resulting from the lack of amenable restriction sites it was decided to bisect the gene at genome position 990348 by introducing an exogenous restriction site. Two primer pairs, shown in table 4.7 were designed to amplify the region for mutagenesis and to introduce a SalI site at position 990348. The first primer pair named 1747mut1F and 1747mut1R was designed to amplify the 1 kb region 989348-990348 of pSymA and to introduce an ApaI site at the 5' end and a SalI site on the 3' end of the PCR product. A PCR was carried out with these primers using *S. meliloti* 2011 genomic DNA as template with the cycling conditions in table 4.8. The resulting PCR product was then purified by gel extraction and subjected to restriction digest by the ApaI and SalI restriction enzymes. In parallel a previously prepared pJQ200sk vector was also digested with ApaI and SalI restriction enzymes. The resulting DNA from these reactions was purified and used in a ligation reaction. Following transformation and screening of transformants a plasmid with a size shift of 1 kb was identified indicating insertion of the PCR product. This was subjected

to restriction analysis to confirm the correct insert. A positive clone was identified and named pJQ1747F1.

Table 4.7: Sequence of the primers used to mutagenize *sma1747*

Name	Sequence
1747mut1F	ATCTCTAGGGCCCTTGGGCTGCACAGCATTT
1747mut1R	ATCTCTAGTCGACGCCACGGTCGAGCCAT
1747mut2F	ATCTCTAGTCGACGGGTCGCTATTTCTTCGG
1747mut2R	ATCTCTAGGATCCCGCGACGTGTTTCTCGTC
47kmscreenF	CGGCTTGAGTTTCAGTTCCAACG
47kmscreenR	CTTTCCACATTGTCGAGACTTGC

Table 4.8: PCR cycling conditions for amplifying *sma1747* mutagenesis fragments

PCR Cycling Conditions	
Annealing Temp	60 °C
Annealing Time	15 sec
Extension Time	20 sec

The second primer pair for the mutagenesis of *sma1747* named 1747mut2F and 1747mut2R, shown in table 4.7 was designed to amplify the 1 kb region 990349-991348 of pSymA and to introduce a SalI site at the 5' end and a BamHI site at the 3' end of the PCR product. A PCR was carried out using these primers on genomic DNA prepared from *S. meliloti* 2011 using the conditions outlined in table 4.8. The resulting PCR product was purified by gel extraction and subjected to restriction digest by SalI and BamHI. In parallel a plasmid preparation of pJQ1747F1 was carried out and subjected to restriction digestion by SalI and BamHI. The resulting DNA was purified and used in a ligation reaction. Following transformation and transformant screening a plasmid with the expected 1 kb size shift was identified. This was subjected to restriction analysis and confirmed as containing the correct 1 kb insert. This vector was named pJQ1747F12 and this is the vector into which the kanamycin resistance cassette will be inserted at the SalI site.

The kanamycin resistance cassette was sourced from the vector pUC4K. A plasmid preparation was carried on the pUC4K vector which was subsequently subjected to a restriction digest by SalI. In parallel the pJQ1747F12 vector was also digested with the SalI restriction enzyme. A ~1.2 kb fragment representing the kanamycin resistance cassette was purified from the pUC4K digest by gel extraction along with the ~7.4 kb fragment representing the linearized pJQ1747F12 vector. These DNA fragments were

then used for a ligation reaction. Following transformation and plating onto LB agar containing kanamycin a plasmid screen was carried out. A correct plasmid was identified by restriction digest and named pJQ1747F12km.

The pJQ1747F12km vector was introduced into *S. meliloti* 2011*rhtX-3* by triparental mating and first recombinants selected for on TY agar containing 1000 μg/ml streptomycin and 80 μg/ml gentamicin. A number of first recombinants were picked and inoculated into 5 ml of TY broth each and allowed to grow to early stationary phase. To select for second recombination events a volume from these cultures was plated onto TY agar containing 5% sucrose and 100 μg/ml kanamycin. Colonies were individually tested for kanamycin resistance and gentamicin sensitivity. A potential mutant was identified, named *S. meliloti* 2011*rhtX-3sma1747* and subjected to a final confirmation by PCR.

Genomic DNA was prepared from a culture of the newly created *S. meliloti* 2011*rhtX-3sma1747*. Using the primers 47kmscreenF and 47kmscreenR in table 4.7 two PCRs were carried out using *S. meliloti* 2011 and *S. meliloti* 2011*rhtX-3sma1747* genomic DNA as template. The PCR products were analysed by agarose gel electrophoresis on a 0.7% gel. The 47kmscreenFR primer pair amplifies the region 990250-990441 of pSymA resulting in a 191 bp fragment in a wild type genome and a 1.4 kb fragment in a correct *S. meliloti* 2011*rhtX-3sma1747* genome. Figure 4.16 is the agarose gel image comparing the wild type result to that of the new mutant. A size shift of 1.2 kb was observed in the mutant representing the kanamycin resistance cassette at the correct position.

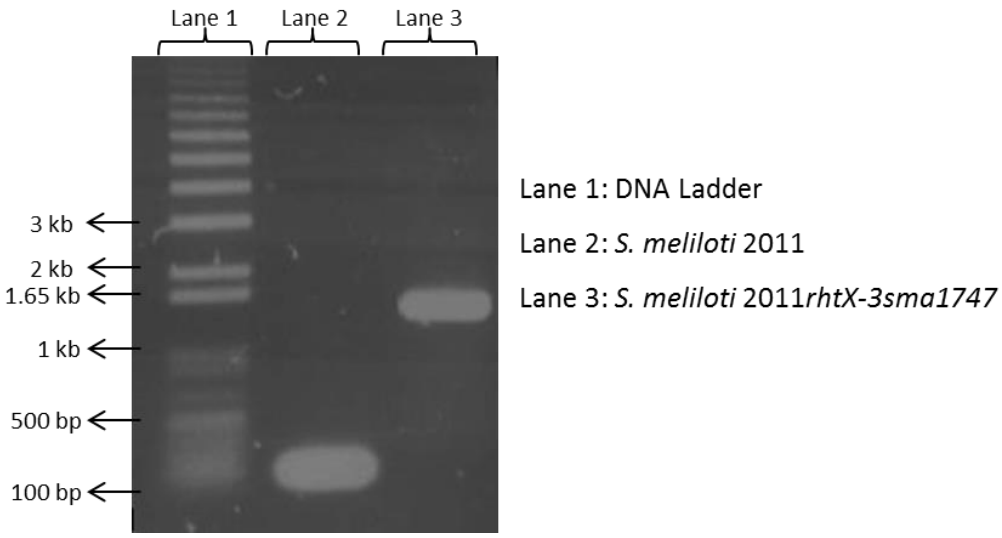


Figure 4.16: Confirmation of the *S. meliloti* 2011*rhtX-3sma1747* mutant by PCR

After confirmation of this mutant the pKC101 vector was introduced by triparental mating to restore the *sma1747* gene *in trans*. The mutant and complemented mutant were then assessed for coprogen utilisation by the iron nutrition bioassay.

4.4.1 Assessment of *S. meliloti* 2011*rhtX-3sma1747* for coprogen utilisation

Iron nutrition bioassays were set up to assess the ability of the new *S. meliloti* 2011*rhtX-3sma1747* mutant and its complement to utilise coprogen. The strain *S. meliloti* 2011*rhtX-3* was included as a positive control as it is the background to the *sma1747* knockout. $FeCl_3$ was included in each bioassay as a positive control to ensure no issues with the inoculum and dH_2O as a negative control as all siderophores are suspended in dH_2O. As described previously 2,2'dipyridyl at 300 μM was included in the media along with IPTG at 500 μM and antibiotics where appropriate. The results of this analysis are shown in table 4.9.

Table 4.9: Assessment of coprogen utilisation by *S. meliloti* 2011*rhtX-3sma1747*

***S. meliloti* Strain**	**Plasmid**	**Coprogen Utilisation**
2011*rhtX-3*	—	+
2011*rhtX-3sma1747*	—	—
2011*rhtX-3sma1747*	pKC101	+

Analysis of the above results confirms that *sma1747* is the outer membrane receptor for coprogen. The knockout mutant of *sma1747* is incapable of utilising coprogen as an iron source. When the *sma1747* gene is present *in trans* on the pKC101 vector utilisation is restored. Utilisation of coprogen by *S. meliloti* 2011*rhtX-3* confirms that the background mutation in *rhtX* is not the cause of the deficient phenotype.

4.5 Antibiotic resistant cassette mutagenesis of inner membrane coprogen transport systems

As discussed previously in order to assess the *sma1746-sma1741* gene cluster for coprogen transport a knockout mutation in *sma1746* was to be created. Based on the organisation of the genes downstream it was likely that this mutant would have a polar effect on *sma1745-sma1741*. The background to this mutant was planned to be *S. meliloti* 2011*rhtX-3*.

In addition to mutating *sma1746* in *S. meliloti 2011rhtX-3* a mutation was to be created in *fhuP* in a background of *S. meliloti* Rm818 to assess its role as a redundant system for coprogen inner membrane transport.

4.5.1 Note on the mutagenesis of *sma1746*

A strategy was designed to introduce a kanamycin resistance cassette at position 988979 that would disrupt *sma1746* at position 4 of its coding sequence. A similar strategy to that used in the knockout of *sma1747* was designed. After carrying out the mutation process numerous possible mutants were identified through resistance to kanamycin and sensitivity to gentamicin however all isolates proved to have the incorrect genotype when assessed by PCR.

After consideration of the failed attempts at mutating *sma1746* it was decided to not continue with the mutagenesis. The desired analysis could by achieved using the *S. meliloti* Rm818*fhuP* mutant complemented in various permutations. A *S. meliloti* Rm818*fhuP* mutant complemented with just pKC101 would represent a genotype identical to a *sma1746* mutant in respect to coprogen utilisation. If the *S. meliloti* Rm818*fhuP* (pKC101) strain proves to be deficient for coprogen utilisation this would show FhuP to be part of a redundant coprogen transport system. If *S. meliloti* Rm818*fhuP* (pKC101) is positive for coprogen utilisation then it would show that the redundant system is located elsewhere on the genome.

Assuming that the *S. meliloti* Rm818*fhuP* (pKC101) is deficient for transport, complementation of this mutant with pKC102 would allow for the activity of *sma1746-sma1741* to be assessed. If *S. meliloti* Rm818*fhuP* (pKC101) is positive for utilisation

then it will be impossible to determine if the *sma1746-sma1741* gene cluster is involved in coprogen utilisation as its role would be occluded by the redundancy.

4.5.2 Mutagenesis of *fhuP* in *S. meliloti* Rm818

As part of the work carried out by O Cuiv et al. (2008) a construct for the mutagenesis of *fhuP* was created using the pJQ200sk+ vector. This mutation was designed to introduce the kanamycin resistance cassette from pUC4K into a naturally occurring SalI restriction site located at position 469 of *fhuP*. This vector was designed and created by Dr. Damien Keogh as part of his PhD thesis and is named pDK2.0K A/S. The vector for complementing *fhuP* was also created by Dr. Damien Keogh. It consists of the entire *fhuP* gene including the ribosome binding site cloned into the pBBR1MCS-5 vector using the ApaI and SacI restriction sites. This vector is named pDK104.

The pDK2.0K A/S vector was introduced into *S. meliloti* Rm818 by triparental mating. Transconjugants were selected for on TY agar containing streptomycin at 750 μg/ml and gentamicin at 80 μg/ml. A number of single colonies were selected from this screen and separately used to inoculate 5 ml aliquots of TY broth. To select for 2nd recombination events a volume was taken from each of these cultures and plated onto TY agar containing sucrose at 5% and kanamycin at 100 μg/ml. Colonies that grew under these conditions were individually examined for resistance to kanamycin and sensitivity to gentamicin. A number of possible mutants were identified, named *S. meliloti* Rm818*fhuP* and subjected to confirmation by PCR.

Table 4.10: Primer sequences for confirmation of *S. meliloti* Rm818*fhuP*

Name	Sequence
FhuPKmF	TTGTCTCGTCTGCTGCTCTGCCGCC
FhuPKmR	TCATTTTCTTCAAGCTTATCCAATAGG

Table 4.11: PCR cycling conditions for confirmation of *S. meliloti* Rm818*fhuP*

PCR Cycling Conditions	
Annealing Temperature	61 °C
Annealing Time	15 sec
Extension Time	90 sec

Two PCRs were set up using the primer pair FhuPKmFR on genomic DNA prepared from *S. meliloti* 2011 and *S. meliloti* Rm818*fhuP* to assess the presence of the kanamycin resistance cassette in *fhuP*. The PCR cycling conditions are shown in table 4.11. The PCR products were analysed by agarose gel electrophoresis on a 0.7% gel. The FhuPKmFR primer pair amplifies the full *fhuP* gene and will give an expected product size of ~1.1 kb for *S. meliloti* 2011 and a ~2.3 kb fragment for a *S. meliloti* Rm818*fhuP* mutant. The results of this analysis are shown in figure 4.17 below. As expected for a positive mutant a size shift of ~1.2 kb was observed indicating the correct phenotype.

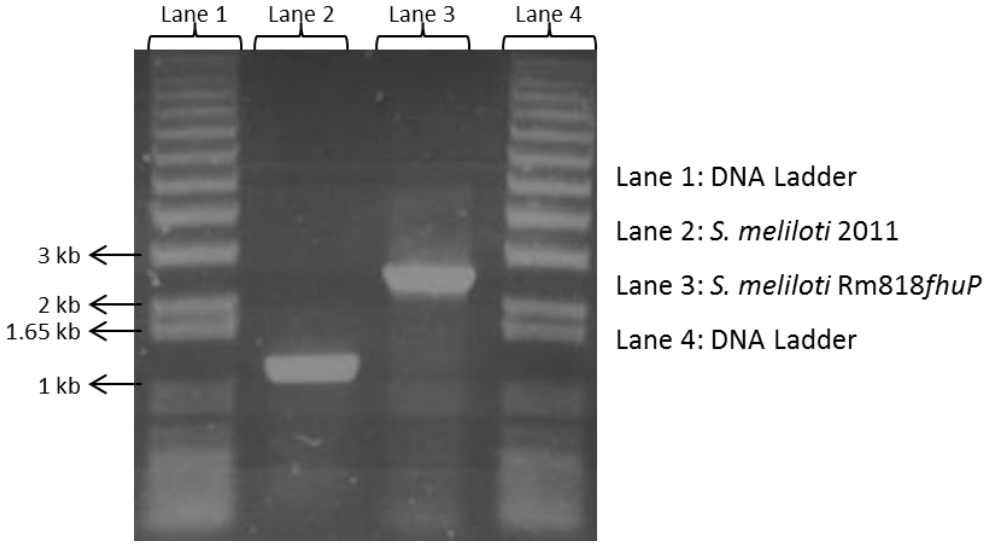


Figure 4.17: Confirmation of the *S. meliloti* Rm818*fhuP* mutant by PCR

After confirmation of the mutant the pKC101 plasmid encoding for *sma1747* was introduced by triparental mating. Once the *S. meliloti* Rm818*fhuP* (pKC101) strain was isolated the pDK104 plasmid was introduced to complement the *fhuP* mutation. The pKC102 plasmid was also introduced into *S. meliloti* Rm818*fhuP* (pKC101) to assess the functionality of the *sma1746-sma1741* gene cluster. These strains were then analysed for coprogen utilisation by iron nutrition bioassay.

4.5.3 Bioassay assessment of the inner membrane transport systems for coprogen utilisation

Table 4.12 below describes the iron nutrition bioassay results obtained from the *S. meliloti* Rm818*fhuP* mutant along with all the various complementations carried out on it. As done previously $FeCl_3$ was included in all bioassays as a positive control and dH_2O as a negative control. A well containing a ferrioxamine B solution was also included as the *fhuP* knockout should result in no utilisation and will act as further confirmation of the *fhuP* mutant having the correct phenotype. As described previously 2,2'dipyridyl at 300 μM was included in the media along with IPTG at 500 μM and antibiotics where appropriate.

Table 4.12: Assessment of coprogen utilisation by *S. meliloit* Rm818*fhuP*

Strain	Plasmid	Ferrioxamine B	Coprogen
S. meliloti* Rm818*fhuP	—	—	—
S. meliloti* Rm818*fhuP	pKC101	—	+
S. meliloti* Rm818*fhuP	pKC101, pDK104	+	+
S. meliloti* Rm818*fhuP	pKC101, pKC102	—	+

The lack of ferrioxamine B utilisation by the *fhuP* negative strains shows that the mutant has the expected phenotype. As the *S. meliloti* Rm818*fhuP* (pKC101) strain still shows coprogen utilisation it demonstrates that FhuP is not involved in coprogen transport. This result leaves the analysis of *S. meliloti* Rm818*fhuP* (pKC101, pDK104) and *S. meliloti* Rm818*fhuP* (pKC101, pKC102) uninformative however the analysis of the former for ferrioxamine B utilisation further demonstrates that the mutant is correct.

This result shows that FhuP is either not a member of the redundant inner membrane transporter complex for coprogen or there is another system present on the genome that is capable of transporting coprogen. There are a number of candidates for the third system if it exists however the practicalities of identifying such a system require a complete redesign of the strategy. As a result of this it was decided against searching for another candidate for inner membrane transport for coprogen.

4.6 Summary and Discussion

As there was still utilisation by the *S. meliloti* Rm818*fhuP* (pKC101) strain this shows that FhuP is not a member of the redundant transport system for coprogen. Whereas as system involving FhuP was the primary candidate for the redundant coprogen transporter it is not entirely unexpected that its mutation does not result in a deficient phenotype. There are a number of possibilities for this:

1. The FhuP protein may not be involved in coprogen transport and the redundant system resides elsewhere in the genome. As the *fhuP* gene is located on the chromosome and the *sma1747* gene on pSymA then it is not surprising they do not overlap in their cognate substrates. However spatial separation of genes did not interfere with the assembly of the "Split System" for the transport of ferrioxamine B, ferrichrome and haem (O Cuiv et al. 2008). This inner membrane system is arranged in two separate regions with one region comprising the *humTUV* genes and the second comprising the *fhuFP* genes. The FhuP protein binds ferrichrome and ferrioxamine B and interacts with *humUV* to internalise the siderophores.
2. The FhuP protein is involved in coprogen uptake and there is a second iron regulated redundant system located elsewhere on the genome. Using the results of the genome wide assessment of iron regulated genes carried out by Chao et al. (2005) a number of possible alternative transport systems can be identified. These are identified as being up regulated in the absence of iron. Table 4.13 is a list of genes encoding inner membrane transport associated proteins up regulated in the absence of iron. In the case of operons all the associated genes are included.

Table 4.13: List of inner membrane transport associated genes up regulated in iron deplete conditions

Gene Designation	Description
sma1860, sma1862-64	ABC transport complex
***rhtX* (pSymA)**	MFS rhizobactin transport
smb21432, smb21431-29	ABC transporter complex
smc00537	MFS transporter
smc00784	fbpA homologue, PBP protein
smc00922	Transport permease
hmuTUV	Haem/hydroxymate siderophore uptake
fhuPF	Hydroxymate siderophore uptake
smc01095	Multidrug efflux pump
smc03971	Multidrug efflux pump

As shown in table 4.13 there are numerous possible transporters expressed under iron deplete conditions in *S. meliloti* 2011 outside of those already characterised. In addition to this list there are numerous genes of unknown function and others with membrane association domains. A number of these transporters can be ruled out as the redundant coprogen transport system as they are found on the pSymA megaplasmid. However, even taking into account the results obtained in the *S. meliloti* Rm818 strain and the analysis of FhuP there are still at least six possible systems that could be involved in coprogen utilisation. As a result mutation of the subsequent candidate genes will require the use of a prohibitive number of varying antibiotic resistance cassettes that would make the screen unpractical to carry out. Random mutagenesis could identify a redundant transporter however utilisation of this method to locate redundant ferrichrome transporters in *P.aeruginosa* PAO1 proved to be fruitless as there were numerous redundant systems (Dr. Paraic Ó'Cuív, Unpublished Results).

3. The FhuP protein is capable of interactin with coprogen and there is a second transport system that is not regulated by cellular iron levels. Such a situation was observed by Letoffe, Delepelaire and Wandersman (2006) when studying haem utilisation in *E. coli* K12. It was observed that in the absence of the primary haem inner membrane transport system homologous to HmuTUV, the dipeptide transport system DppBCDF along with the dipeptide binding periplasmic binding proteins DppA and MppA facilitated haem transport across the inner membrane. Whereas it is unlikely that the DppABCDF system plays any role in coprogen utilisation there may be a similar phenomenon occurring in that a transport system historically unrelated to siderophore/iron uptake may be partially compensating transport in the absence of the primary transporters.

In addition to identifying redundancy at the inner membrane for coprogen utilisation the cognate outer membrane transporter was identified as Sma1747. This gene was identified by two methods; initially by complementation of the *S. meliloti* Rm818 strain with the outer membrane receptor and secondly by creation of a kanamycin cassette disruption of the *sma1747* gene in an *rhtX* background and subsequent restoration of the phenotype by complementation. Analysis of coprogen utilisation in *E. coli* K12 undertaken by Hantke (1983) identified the cognate outer membrane transporter and

designated it as FhuE for **F**erric **h**ydroxamate **u**ptake E and it is proposed that *sma1747* is reassigned as *fhuE* based on this.

In light of this the *S. meliloti* 2011 genome was analysed by BLASTp using FhuE as the query sequence. The top three homologies are given in table 4.14.

Table 4.14: *S. meliloti* 2011 proteins similar to FhuE from *E. coli* K12 by BLASTp

Protein	Genome Location	Query Coverage (%)	Identity (%)	E Value
FhuA	Chromosome	96	25	3e^-33
Sma1747	pSymA	96	23	3e^-26
FoxA	Chromosome	93	22	5e^-18

These homologies show the same pattern as observed in table 4.1 as the outer membrane receptors for ferrichrome, ferrioxamine B and coprogen are very similar to one another.

4.7 Conclusion

The phenomenon of coprogen utilisation fits into the broader category of fungal siderophore pirating by bacteria. Focusing on this phenomenon in *S. meliloti* 2011 it has been shown by the analysis presented here and by previous studies carried out by O Cuiv et al. (2008) that *S. meliloti* 2011 can utilise coprogen, ferrichrome and ferrioxamine B all of which are produced from fungal sources. Fungal siderophores sources have been discussed as part of a recent review Hider and Kong (2010) in which the most prominent fungal siderophore producers were discussed. The siderophores produced by these organisms were found to be primarily ferrichrome and coprogen or a derivative of one or the other. Studies into the prevalence of fungal siderophores have found that in some regions soil siderophore levels, specifically ferrichrome and the closely related ferricrocin can reach concentrations of 10 nM (Essen et al. 2006; Prabhu, Biolchini and Boyer 1996). This is not surprising as the known producers of these siderophores are closely associated with plants such as the plant pathogen *Ustilago maydis* the causative agent of corn smut which is a prominent example. Of greater interest is that members of the mycorrhizal fungi have been shown to produce siderophores related to the hydroxamate ferrichrome structure (Winterberg et al. 2010; Prabhu, Biolchini and Boyer 1996; Mei, Budde and Leong 1993). If siderophore production is a common feature of members of the mycorrhizal fungi then it is not surprising that such high titres of hydroxamate fungal siderophores are found in soil

samples. This is due to the near ubiquitous occurrence of symbiotic relationships between mycorrhizal fungi with all but a few plant species. The primary form of fungal symbiosis is facilitated through intracellular structures known as arbuscules. These structures allow for nutrient exchange to occur between the plant and symbiotic fungus and are formed specifically by members of the arbuscular mycorrhiza. These relationships are so widespread and common that it is thought to represent the preferred method of mineral uptake in plants (Gutjahr and Parniske 2013).

If siderophore production is common in members of the arbuscular mycorrhiza then this would provide an interesting hypothesis for xenosiderophore utilisation in *S. meliloti*. It is known that *S. meliloti* forms a close intracellular relationship with the plant species *Medicago sativa* commonly known as alfalfa. The outcome of this symbiosis is the formation of an intracellular bacteroid that fixes atmospheric nitrogen by reduction to ammonia which enhances the plant health and growth. The bacteria benefit as they are supplied with a direct carbon source in ample quantities (Capela et al. 2006). In addition to the bacterial symbiosis formed with *S. meliloti, Medicago sativa* forms another close symbiosis with the arbuscular mycorrhizal fungus *Glomus intraradices*. This *Glomus-Medicago* symbiosis results in the plant gaining valuable access to insoluble phosphorus along with other minerals and also benefits the fungus through the plant feeding the fungus a carbon source. A study carried out which monitored plant growth and health after inoculation with each symbiont in singular and dual inoculation in the presence of an insoluble phosphorus source, only obtainable by the fungus, yielded interesting results. As expected each symbiont increased the biomass of the plant in relation to the un-inoculated control and an additive effect was observed in the dual inoculated plant (Stancheva I. et al. 2008). Studies into the life cycles of both these symbionts displayed a remarkable level of similarity. In brief the symbiont must first exit the plantonic lifestyle that is usually initiated by an exudate produced from the plant roots. For the fungus this represents spore germination as it is an obligate symbiont. The cells then migrate into the plant and colonise the plant in the extracellular space. The intracellular symbiosis is then formed with the formation of arbuscles in the fungal symbiosis and the formation of bacteroids in the bacterial symbiosis. Throughout this process the level of nutrient available to the symbionts is greatly increased which results in accelerated growth. After the symbiosis each symbiont is released from the plant and returns to the soil (Denison and Kiers 2011). The process of symbiosis has been examined in detail

for each organism but the activities of the symbionts in a free living state has not been explored in great detail.

To return to iron acquisition in *S. meliloti* this tri-species interaction may involve more than just the two symbionts interacting with the plant partner but rather a more complex interspecies interaction. Members of *Glomus* species have been shown to produce a siderophore but the structure of this siderophore or siderophores is unknown (Cress, Johnson and Barton 1986). If *S. meliloti* was capable of utilising a siderophore produced by *Glomus intraradices* this would relieve the iron stress from *S. meliloti* and promote growth in a free living environment which would increase the likelihood of a symbiosis. However for such a situation to be proven the structure of the siderophore produced from G*lomus intraradices* would have to elucidated and subsequently purified to assess its ability in promoting growth of *S. meliloti*. Even in the absence of such a relationship the ability of *S. meliloti* to acquire fungal siderophores provides it with a readily available iron source that can be acquired without the metabolic cost of producing its own siderophore rhizobactin 1021, which results in an improvement of the overall fitness of the organism.

Chapter Five

Siderophore Mediated Regulation of Outer Membrane Receptor Expression

5.1 Introduction

Acquisition of iron is essential to the survival of *S. meliloti* 2011. The two most prominent methods by which this is accomplished have been explored in detail in the previous chapters which are production of rhizobactin 1021 and siderophore pirating at the expense of other organisms. While these systems are vital to *S. meliloti* in iron deprived situations they can also result in cell toxicity if incorrectly regulated. Iron toxicity presents a major problem to a cell as an overabundance of iron leads to high levels of oxidative stress and free radical formation resulting in cell damage.

To prevent the accumulation of toxic levels of iron, *S. meliloti* primarily relies on the activity of the global gene expression regulator RirA. The regulon of the RirA protein was assessed by microarray analysis by Chao et al. (2005). This study determined that RirA acted in a global manner to repress numerous genes involved in iron transport, energy metabolism and exopolysaccharide production in iron replete conditions. This was shown by comparison of genes differentially expressed in a *rirA* mutant to genes differentially expressed in iron deplete conditions. It was determined that the *rirA* mutation did not account for all changes in iron deplete conditions but it did account for all genes putatively annotated as encoding iron transport functions that were also expressed to a higher level in iron deplete conditions. However, it is worth noting that the genes *fhuA* and the newly described *fhuE* encoding for ferrichrome and coprogen transport respectively were not differentially expressed in either the *rirA* mutant or iron deplete conditions.

A second more specific study into the role of RirA in *S. meliloti* focused on how the protein affected the rhizobactin 1021 regulon (Viguier et al. 2005). The results obtained in this study correlated with the observations made in the microarray analysis by Chao et al. (2005) as the rhizobactin 1021 regulon was expressed at a higher level in a *rirA* mutant. In addition to RirA, a second transcriptional regulator has been shown to influence the expression of the rhizobactin 1021 biosynthesis and transport genes namely, the AraC-like protein RhrA (Lynch et al. 2001). This protein was identified in the initial description of the rhizobactin 1021 regulon and a knockout mutant resulted in abolition of siderophore production. It was also found through RNase protection experiments that RhrA induced the expression of *rhbA* and *rhtA* as these transcripts were not detected in an *rhrA* mutant. These results were confirmed by Chao et al.

(2005) as they assessed a *rirA-rhrA* double mutant. The single mutant of *rirA* resulted in constitutive siderophore production and the single mutant in *rhrA* resulted in siderophore abolition as observed by (Lynch et al. 2001). The phenotype of the *rirA-rhrA* double mutant displayed no siderophore production. This is significant as it demonstrates that the expression of the rhizobactin 1021 regulon is dependent on the activity of the AraC-like protein RhrA which is in itself repressed by RirA under iron replete conditions. This results in an enhanced local activation of gene expression that is directly linked to the iron status of the cell.

With this example of a global regulator, RirA, affecting change via the local activator RhrA, it is worth investigating whether such a phenomenon is occurring with regard to *fhuE* and *fhuA*. Analysis of the genome location of these two genes shows that each is located adjacent to a putative AraC-like regulator which raises the possibility of local regulation mediated via *sma1749* for *fhuE* and *smc01610* for *fhuA*. The gene arrangements are given in figure 5.1.

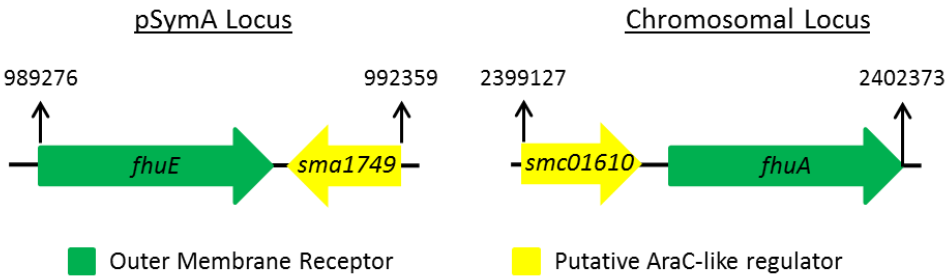


Figure 5.1: Genome position and gene arrangement of *fhuE, sma1749, fhuA* and *smc01610*

The gene map shows that *sma1749* is positioned proximal to and converges with *fhuE*. The *smc01610* gene is positioned proximal to *fhuA* with both genes transcribed from the positive strand. A previous study carried out by Keogh (2008) showed that *smc01610* and *fhuA* appear to also be co-transcribed or that the specific gene arrangement is essential for correct function. It was found that a *smc01610* knockout by insertional inactivation resulted in the loss of the ability to utilise ferrichrome. When *smc01610* was present *in trans* this did not restore the ability to utilise ferrichrome. The co-location of the two outer membrane receptors with an AraC-like protein taken with the lack of a response in the RirA microarray study allows for the possibility that the outer membrane receptors are regulated at a local level.

To assess the possibility of a local regulation of the outer membrane receptors for ferrichrome and coprogen the expression levels of these genes will be monitored by reverse transcription quantitative PCR. Each of the two genome loci will be monitored in iron replete and iron deplete conditions both in the presence and absence of the cognate siderophore.

Prior to analysing the expression of each system the properties of both of the putative AraC-like proteins, Sma1749 and Smc01610 will be analysed in greater detail.

5.2 Bioinformatic analysis of Sma1749 and Smc01610

The vast majority of AraC protein family members act to activate expression of their target genes. They are generally encoded as ~300 amino acid proteins that have two distinct functional domains. The characterising feature of the AraC family is the presence of a helix-turn-helix region present at the C-terminus of the protein which covers ~100 amino acids. This region is responsible for recognition and binding of specific DNA sequences associated with the promoter region of the target gene. This region is conserved among all members of the family. The second domain of AraC proteins is at the N-terminal and is more variable than the C-terminal domain. This region is responsible for substrate recognition and also allows the formation of homodimers. The two domains act in unison with the N-terminal domain acting as a sensor for the cognate substrate and the C-terminal acting to bind DNA and induce expression of the target gene (Gallegos et al. 1997).

5.2.1 Analysis of Sma1749

The protein Sma1749 is predicted to be 299 amino acids in length with a molecular weight of 33 kDa and a pI of 9.35. The protein sequence is given in the previous chapter as figure 4.6 along with a brief analysis by the Interpro program but will be discussed with regard to a possible regulatory role exerted on *sma1747*. The results of the Interpro analysis did not classify Sma1749 as a member of the AraC-like family of transcriptional regulators. However, an AraC-type helix-turn-helix structure was identified which comprises the region 190-294 at the C-terminal. As the N-terminal of AraC-like proteins is variable due to its function in substrate binding this region displays no protein signatures that would assign it to a specific family or function.

5.2.2 Analysis of Smc01610

The Smc01610 protein is 294 amino acids in length with a molecular weight of 33 kDa and a pI of 8.99. The amino acid sequence for Smc01610 is given in figure 5.2.

```
MTFQPRMQNRISGFSIIGGLNRREWNGVVADVWDVECVPHAGGYYVAEDP
RMFIVLDARGGGNCRVKLAANGKGAVQNYHRQALSYIPAGMELWTDVVDI
HYIRHLDLHFDVDALGRRLKEDLDAAAIETPRLMFQDERFLTLAGLIAAE
CLNPQPLHDLYGDSLTVALFIDLMKIGKRSGRKRSQLAAWQLRRAVDFIE
ENFARNVRLEELAGLTGLSQSHFSHAFKASTGVAPHQWHMNARVERAKQM
LLRSDAPLTSIAAETGFADQAHFTRVFRKAVGTTPALWKKSHTA
```

Figure 5.2: Protein sequence of Smc01610

The Interpro analysis of Smc01610 places it in the protein family of AraC-type transcriptional regulators. Similar to Sma1749 there is a clear C-terminal domain that spans position 188-292. This domain forms the helix-turn-helix structure that is essential for the DNA binding activity of the protein. Also the N-terminal region displays no identity that is identifiable with a specific protein family which is in keeping with the structure of other AraC-like proteins.

5.2.3 Summary of AraC-type protein analysis

Both of the putative AraC-like proteins, Sma1749 and Smc01610 show similarities to the AraC protein family. It is interesting that the two regions display different gene arrangements. The differing position of AraC-like genes with that of the gene that they regulate may suggest that they influence their target gene through two different mechanisms. The following qPCR analysis gives insights into the possible influence the AraC-like proteins Sma1749 and Smc01610 exert on the genes encoding outer membrane receptors *sma1749* and *smc01610.*

5.3 Background to coprogen and ferrichrome transport gene expression analysis

5.3.1 Background to qPCR

RT-qPCR is a powerful tool for gene expression analysis. It allows for minute changes in gene expression to be measured relatively quickly and reliably. There are numerous methods that allow for DNA amplification to be monitored by quantitative PCR with the two most prominent being fluorescent hydrolysis probes or the intercalating dye Sybr Green. There are numerous advantages and disadvantages to each method with the Sybr Green method being of benefit as it is much less expensive than the use of hydrolysis probes. Also Sybr Green allows for issues with the qPCR assay to be identified as it detects non-specific products and primer dimers, whereas hydrolysis probes only identify the target product which can mask underlying reaction issues. Hydrolysis probes allow for only the correct PCR product formation to be monitored which is beneficial in multiplex situations as it allows multiple assays to be run in one tube. However, regardless of the detection method used, the formation of secondary PCR products must be as low as possible. Secondary PCR product formation results in primers and polymerase not being utilised efficiently which can negatively impact DNA detection if the unwanted products are prominent. As a result of this it is important to optimise and check each assay prior to use to ensure the correct product is formed and that no products are formed in the no template controls. The Sybr Green method of tracking DNA amplification was used in this analysis (Bustin et al. 2009).

The measurement of gene expression change was achieved by normalising gene expression to a gene unaffected by the variations in conditions used and is referred to as a reference gene or a housekeeping gene. Once gene expression is normalised the fold expression change can be determined. The reference gene used for normalisation in these analyses was *smc03979,* also known as *gapA* which encodes for a probable glyceraldehyde 3-phosphate dehydrogenase involved in the metabolism of glucose. This gene has been used in previous iron related gene expression experiments in *S. meliloti* Amarelle et al. (2010) and yielded reliable results. Also *smc03979* was not found to be differentially expressed in the microarray study carried by Chao et al. (2005) in which expression changes were monitored between iron replete and iron deplete cultures.

Using the values obtained for the *gapA* gene the expression change in the genes of interest was determined by the delta-delta Ct method (Livak and Schmittgen 2001).

The choice of reference gene in RT-qPCR is vital if accurate results are to be obtained and perhaps equally important to this is that each qPCR assay is preforming optimally. To achieve this, each primer set is designed to amplify a region of between 70 bp and 150 bp in length. A short amplicon is desirable as it ensures that the PCR product is fully extended at the end of each cycle and also that extension occurs very quickly which helps prevent the formation of non-specific products and primer dimers. Each primer pair must anneal at 60 °C as the qPCR assays will be run concurrently and the annealing temperatures cannot by varied across the samples. As the annealing temperature cannot be altered the occurrence of unwanted PCR products is avoided by incrementally optimising the primer concentration used in each assay. The primer concentration is optimised to the point at which only one PCR product is attained in the template containing sample and no non-specific PCR products are formed in a no template control (Bustin et al. 2009).

In addition to each qPCR assay performing optimally from a technical viewpoint the integrity of the samples to be analysed is also vital. Care was taken to harvest cultures at an OD_{600} value representative of early-mid exponential phase as cells at this point of their growth cycle would be healthy and should not be under any nutrient stress other than the intended iron stress. In addition to harvesting cells at the correct growth phase the extracted RNA was assessed for the presence of contaminating protein, salt and solvents. This was done using a Nanodrop 1100 Spectrophotometer that analysed each RNA extract at wavelengths of 230 nm, 260 nm and 280 nm. The 260/280 ratio represents the level of protein contamination in the sample as the 280 nm wavelength measures protein levels and the 260 nm wavelength measures RNA levels. The 260/230 ratio represents the level of salt/solvent contamination in the sample as the 230 nm wavelength detects a wide range of dissolved chemical contaminants in the sample. A highly pure preparation of RNA would have a 260/280 and 260/230 ratio of between 1.8 and 2. A DNAse digest is carried out on each sample to ensure that no contaminating genomic DNA is present in the RNA preparation.

5.3.2 Optimisation of qPCR assays

As mentioned earlier it is essential to design and optimise each qPCR assay so that each assay amplifies a product of the optimal size and forms no non-specific products. Primers were initially designed using the realtime PCR Tool provided online by IDT DNA Technologies, http://eu.idtdna.com/scitools/Applications/RealTimePCR/. The optimal primer concentration was determined by incrementally varying the primer concentration for each assay. The range of optimisation is generally between a final reaction concentration of 1 μM and 100 nM. Each assay is tested at a number of primer dilutions and the highest primer concentration that results in a single PCR product with no amplification in the no template control being selected for use in the analysis of gene expression. The primers used to monitor gene expression of each gene are given in table 5.1 and the optimised detail of each assay is given in table 5.2.

Table 5.1: List of primers used in to monitor the gene expression of *fhuA*, *fhuE* and associated genes

Primer Name	Sequence	Target Gene
Smc03979 qFor2	GATCCACTCCTACACCAACG	*gapA*
Smc03979 qRev2	AGCTTGCCCTTGAGTTCC	
FhuA qFor1	TCCAGATAACCAAGCAGAACG	*fhuA*
FhuA qFor2	GGTATAGGAGAAGGCGCTTATG	
Smc01610 qFor1	CGCATGCAGAACAGGATCAG	*smc01610*
Smc01610 qRev1	AACCACACCATTCCACTCTC	
Sma1747 qFor1	GGGTAAGCAGTATGAAGTCGG	*fhuE*
Sma1747 qRev1	ATGTCTGACCTTTTCGACGG	
Sma1749 qFor1	TCATAAGGATCGGGTGTTGC	*sma1749*
Sma1749 qRev1	TGGTAGGGCTCTTTATTTCGTC	
FhuP qFor2	GGTCTTCGAGAGAGAAGAACAG	*fhuP*
FhuP qRev2	GCTTCCATGAGAACTTTCGG	

Table 5.2: Expected product size and primer concentrations used for each qPCR assay

Target Gene	Amplicon Size	Optimal Primer Concentration
gapA	150 bp	333 nM
fhuA	149 bp	500 nM
smc01610	71 bp	333 nM
fhuE	72 bp	333 nM
Sma1749	148 bp	333 nM
fhuP	119 bp	500 nM

Upon optimisation of each assay the gene expression could be monitored for each gene of interest.

5.4 Analysis of the role of coprogen in the expression of *fhuE* and *sma1749*

A single colony of *S. meliloti* was picked from a fresh streak plate and used to inoculate a 5 ml TY broth. This culture was grown at 30 °C at 200 rpm for 2 days to allow the culture to reach late exponential/early stationary phase. Once the culture achieved a high cell density it was aseptically diluted to an absorbance of 1 at OD_{600} and used as a 1:100 inoculum into 25 ml of TY broth. The TY broth was supplemented accordingly to stimulate iron stress condition in the presence and absence of coprogen. Each culture was set up in triplicate to ensure biological significance of the analysis. The conditions used are outlined in table 5.3 below.

Table 5.3: Growth conditions for expression analysis of *fhuE* and *sma1749*

Conditions Required	Supplement Added	Supplement Concentration
Iron Replete	None	N/A
Iron Deplete	2,2'Dipyridyl	300 µM
Iron Replete with Coprogen	Coprogen	1 µM
Iron Deplete with Coprogen	2,2'Dipyridyl Coprogen	300 µM 1 µM

The concentration of 2,2'dipyridyl and coprogen used were optimised as part of a previous study (Keogh 2008). The 2,2'dipyridyl concentration was found to be sufficient to result in a noticeable decrease in growth but still allow growth to proceed. It was found that the addition of coprogen to an iron deplete culture restored the growth rate to levels similar to the replete culture. The addition of coprogen to the iron replete culture did not alter the growth rate.

Growth was allowed to proceed until an OD_{600} of ~ 0.6 was reached, at which time the culture was harvested and the RNA extracted. A 15 ml aliquot of the culture was removed and centrifuged at 4,000 rpm for 5 min. The supernatant was removed and the cells suspended in 1 ml of PBS and subsequently transferred to a 1.5 ml reaction tube. The cells were centrifuged at 6,000 rpm for 2 min to collect the cells and the supernatant removed. The cell pellet was dissolved in 1 ml of Tri-reagent and the RNA extracted by the procedure described in section 2.5.4. Upon completion of the extraction the RNA sample was DNase treated to remove contaminating genomic DNA and a cDNA synthesis reaction was carried out as described in section 2.5.5. In brief, the DNase treated RNA was normalised to a concentration of 100 ng/µl in RNase free water

and 10 µl of this added to 10 µl of a 2X mastermix of the reverse transcriptase, random hexamer primers, dNTPs and the reaction buffer. The reaction was allowed to run to completion and the resultant cDNA used to analyse the expression of *fhuE* and its possible regulator *sma1749*.

Using cDNA prepared from each of the conditions in table 5.1 a qPCR plate was set up to monitor the expression of the genes of interest. The gene monitored as part of this study were as follows; *gapA* acted as the reference gene against which the expression change of each other gene could be determined, *fhuE* is the first gene of interest as it encodes the outer membrane receptor for coprogen, *sma1749* is the second gene of interest as it possibly exerts influence over *fhuE* expression and *fhuP* was included as a control gene of interest. Expression of *fhuP* was found to be controlled by RirA by Chao et al. (2005) and served as an internal reference off which to judge the iron status of the cell. The lightcycler conditions used for each assay are as follows;

Stage 1: Step 1: 95 °C for 10 min

Stage 2: Step 1: 95 °C for 15 sec
Step 2: 60 °C for 15 sec
Step 3: 72 °C for 37 sec
} 40 Cycles

Stage 3: Step 1: Melt Curve Analysis

Stage 1 was carried out to ensure complete denaturation of the cDNA at the beginning of the reaction. Acquisition of the dye fluorescence was carried out in step 3 of the second stage. Stage 3 was a melt curve analysis of the DNA product produced in each tube. This was used to troubleshoot anomalous values attained in some reactions especially during assay optimisation.

As mentioned earlier the expression change of each gene was determined by the delta-delta Ct method. This method requires the use of a reference gene, *gapA* in this analysis, along with a control condition which the expression change can be compared. In all the experiments the iron replete culture was used as the control condition. With this in mind the following calculation was employed to determine the fold expression increase or decrease in the test culture in comparison to the control iron replete culture.

1. **ΔCt = Ct of gene of interest – Ct of reference gene**
2. **ΔΔCt = ΔCt of test culture – ΔCt of control culture**
3. **Fold expression change = $2^{-\Delta\Delta Ct}$**

In equation 1 the ΔCt is calculated for each gene of interest in both the test culture and the control culture which results in the two values of ΔCt of test culture and ΔCt of control culture. In equation 2 the difference between these two values is determined, ΔΔCt, by subtracting the ΔCt of the control culture from the ΔCt of the test culture. This value is utilised in equation 3 as the value 2 to the power of –ΔΔCt, which results in numerical value for the fold change in expression. A value greater than 1 donates an up regulation and a value less than 1 donates a down regulation.

Using this methodology the effect of coprogen on the expression of *fhuE* and *sma1749* was assessed in both iron replete and iron deplete environments.

In addition to qPCR reactions that assessed the expression profile of the genes of interest two types of controls were included in each analysis namely a no-RT control and a no-template complete. The no-RT control tested against genomic DNA contamination of the RNA sample and the no template control tested against DNA contamination of the reaction constituents.

5.4.1 Expression analysis of *fhuE* and *sma1749*

The expression change of the three genes of interest, *fhuE*, *sma1749* and *fhuP* was monitored in iron deplete conditions and compared to the control iron replete conditions. This analysis will determine which genes are directly under the influence of the iron status of the cell, most likely through the activity of RirA. The results of this analysis are given in figure 5.3.

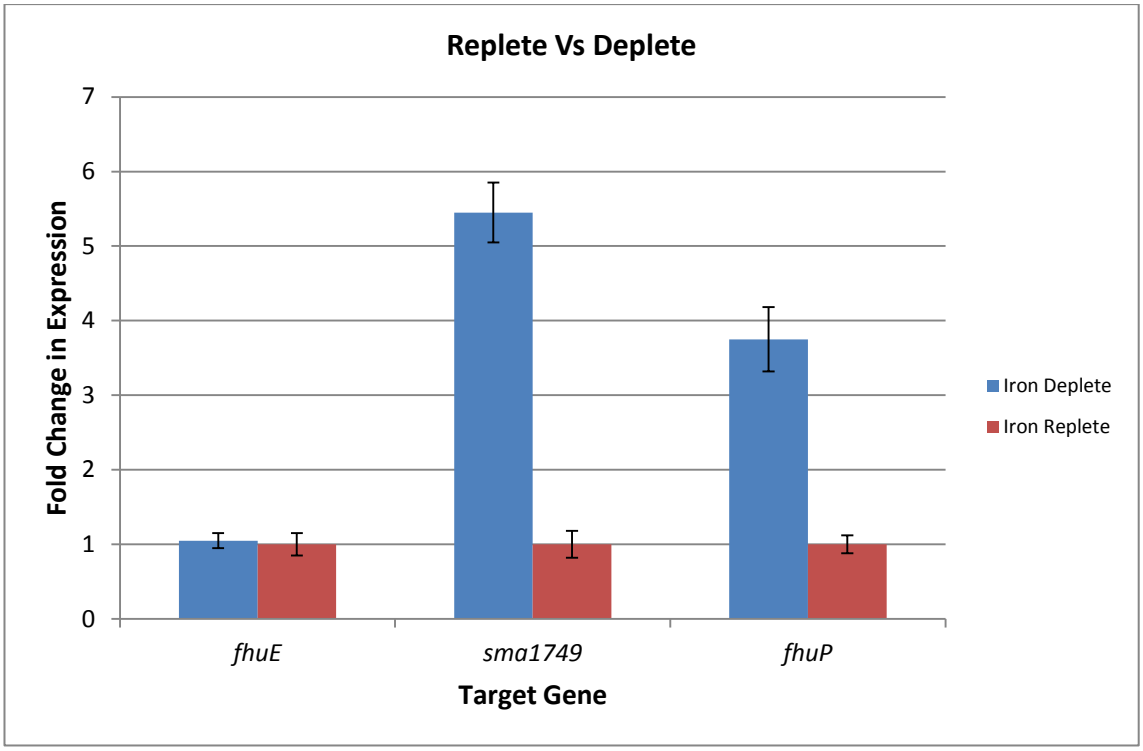


Figure 5.3: Expression analysis of *fhuE* and *sma1749* in iron deplete conditions

To show the expression change clearly the iron replete sample was set to an arbitrary value of 1 and the expression change in the iron deplete sample presented as a fold change off this value. As can be seen there is no response from the *fhuE* gene to the low iron availability. In contrast to this the *sma1749* demonstrates a ~5.5 fold increase in expression in response to a low iron environment. The increase in *fhuP* expression indicates the conditions of the experiment were correct and the culture was under iron stress in the iron deplete sample.

These results indicate that the genes *sma1749* and *fhuP* are affected by the iron status of the cell, which may by through the activity of RirA. The stability of expression of *fhuE* shows that iron limitation is not sufficient to initiate gene expression.

In light of the lack of induction of *fhuE* in an iron deplete environment the effect of the presence of coprogen in a deplete environment was assessed to determine if coprogen is required for *fhuE* expression. Figure 5.4 shows the results obtained when gene expression of the three genes of interest was compared between iron replete conditions and iron deplete conditions in the presence of 1 µM coprogen.

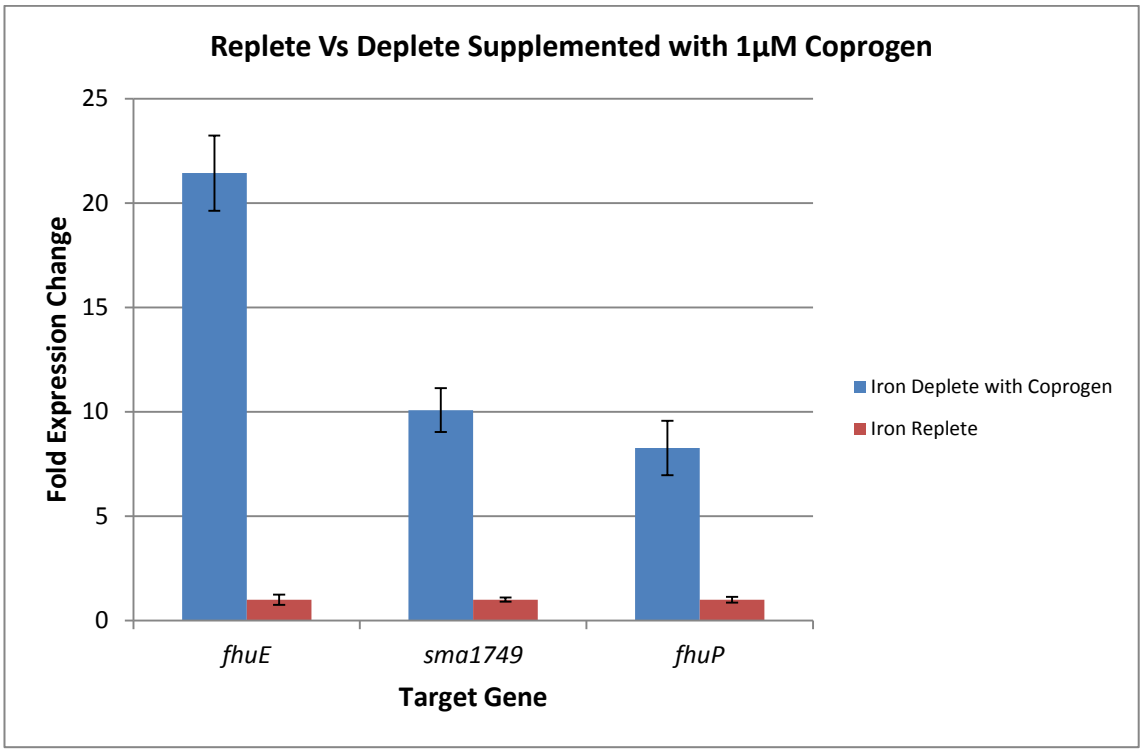


Figure 5.4: Expression analysis of *fhuE* and *sma1749* in iron deplete conditions supplemented with 1µM coprogen

As before the iron replete sample was set to a value of 1 to allow for direct comparison of the fold increase in gene expression. The addition of coprogen to the iron deplete sample has resulted in a drastic increase in the expression of *fhuE*. The expression of both *sma1749* and *fhuP* show the same trend as that observed in the iron deplete sample which indicates that they are induced by iron limited conditions. It is of interest that *fhuP* is induced to a greater level than observed in the iron deplete sample which could indicate that the siderophore is playing a role in its expression. The key difference in the comparison of iron deplete and iron deplete in the presence of 1 µM coprogen to iron replete is the drastic increase in expression of *fhuE*. This result indicates that the *fhuE* gene is most likely repressed until coprogen is present, at which point the receptor is strongly expressed.

The comparison between figures 5.3 and figure 5.4 shows that induction of *fhuE* is dependent on the extracellular environment being low in iron and coprogen being present. However this result does not show whether this effect is solely due to the presence of coprogen. If *fhuE* expression responds to coprogen alone it would demonstrate a system independent of the iron stress on the cell. If *fhuE* is not induced in the presence of coprogen in a replete sample it would prove that iron limitation must play a role in controlling the *sma1749/fhuE* system.

To determine if the presence of coprogen alone can induce *fhuE*, the expression change was monitored in an iron replete culture supplemented with 1 µM coprogen. The results of this analysis are given in figure 5.5.

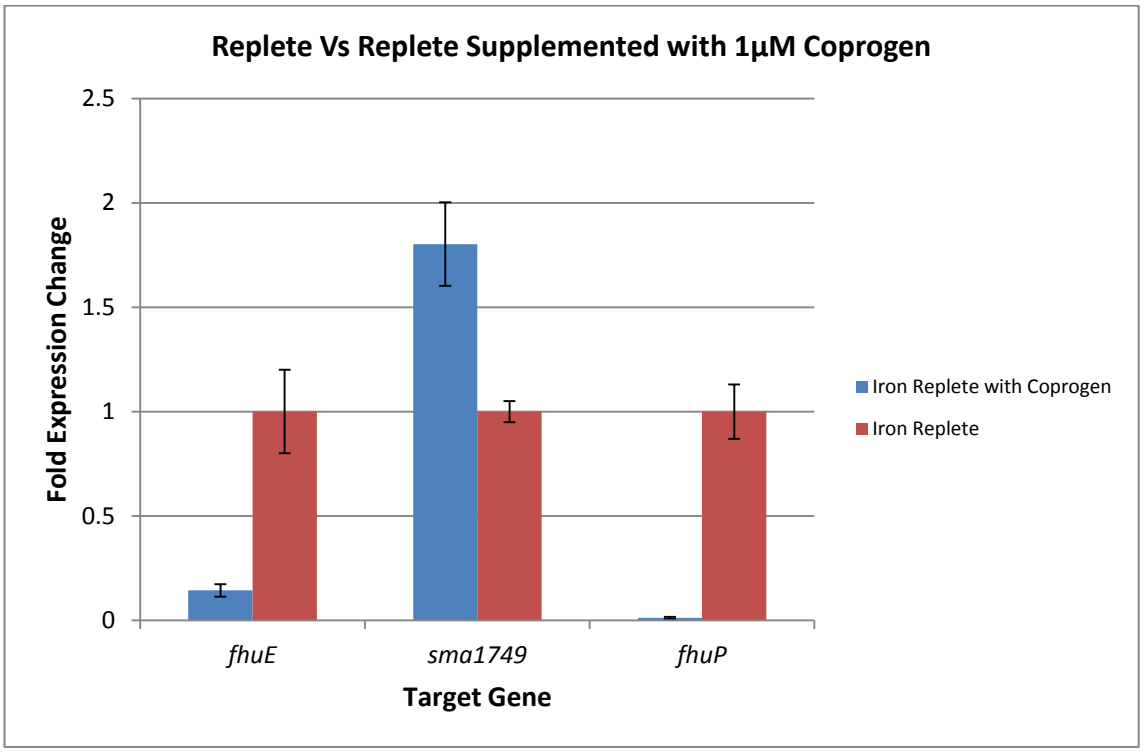


Figure 5.5: Expression analysis of *fhuE* and *sma1749* in iron replete conditions supplemented with 1 µM coprogen

As before the expression levels of the iron replete sample was set to 1 to demonstrate the fold expression increase or decrease. These results confirm that the induction of *fhuE* is dependent on both low iron conditions and the presence of coprogen. The down regulation of *fhuE* suggests that the induction mechanism of this system relies on the cell to be iron deprived for activation. It appears that the basal level of *fhuE* allows for enough FhuE to be present in the outer membrane to essentially feed the cell coprogen. This is evident by the significant drop in expression in *fhuP* in the presence of coprogen

which indicates that the gene is repressed more strongly than the normal iron replete conditions. The up regulation of *sma1749* goes against the observed results in figure 5.3 where it appears to be responsive to the iron levels in the cell. This increase in the presence of coprogen indicates that coprogen can influence the expression of *sma1749* along with *fhuE*. Such a system where the presence of the effector molecule, coprogen, induces the expression of the regulatory gene, *sma1749*, will allow for rapid induction of *fhuE* if conditions were the change from iron replete to iron deplete but also ensure that *fhuE* expression only occurs when the cell requires iron. Further experimentation would be required to show this which was not covered as part of this analysis.

5.4.2 Summary of the expression analysis of fhuE and sma1749

The expression analysis of the outer membrane receptor *fhuE* and its associated AraC-like gene, *sma1749* presented here demonstrates that the induction of *fhuE* is dependent on the presence of coprogen and most likely the activity of RirA. This was demonstrated through the ~20 fold increase in *fhuE* expression in the presence of coprogen in an iron deplete environment. This is in contrast to the profile observed in the iron deplete and iron replete conditions supplemented with coprogen. Iron deprivation did not change the expression levels of *fhuE* and surprisingly *fhuE* was down regulated when coprogen is present in a replete environment. These results taken together suggest that perhaps RirA is acting through a second regulatory functional that acts in the presence of coprogen. The prime candidate for this role would be Sma1749.

As shown previously *sma1749* was induced in conditions of iron stress. This would suggest that RirA is repressing *sma1749* when there is no iron stress on the cell. Also, *sma1749* was expressed to a higher level in an iron replete environment in the presence of coprogen, which suggests that the presence of coprogen can also induce *sma1749*. The assay for *fhuP* was included as a control, however the greater induction observed in the coprogen deplete sample compared to the deplete sample suggests that the presence of the siderophore may also be playing a role here. However as it is not up regulated in a replete environment in the presence of coprogen it indicates that iron deprivation is the primary control condition. As this analysis is based solely on gene expression levels, the activity of Sma1749 can only be speculated on. However assuming that Sma1749 interacts with coprogen it is possible to present a possible benefit to such a situation of

dual responsiveness. As *sma1749* responds to both iron deprivation and the presence of coprogen, this would position the cell so that it could respond rapidly if conditions were to change from one where one stimulant is present to one with both present. This would allow the cell to immediately benefit from the presence of coprogen as an iron source but only if required.

The inclusion of an assay to monitor the expression of *fhuP* proved valuable as it acted as an indirect method of assessing the iron homeostatic position of the cell.

These results will be used to suggest a possible model for the activity of Sma1749 and how it induces *fhuE* in the correct conditions. Also possible binding sites for Sma1749 protein will be discussed as part of this model. This will be presented after the expression analysis of *fhuA* and *smc01610.*

5.5 Expression analysis of *fhuA* and *smc01610*

A similar analysis as that carried out on the coprogen associated genes was carried out on the genes *fhuA* and *smc01610*.

A single colony was picked from a fresh streak plate of *S. meliloti* and inoculated into 5 ml of TY broth. This culture was incubated at 30 °C at 200 rpm for two days until late exponential phase/early stationary phase was reached. The culture was diluted to an OD_{600} reading of 1 and used to inoculate 25 ml of TY broth as a 1:100 inoculum. The TY broth was inoculated with various combinations of ferrichrome and 2,2'dipyridyl to induce the desired conditions of interest. Each condition was inoculated in triplicate to ensure the biological significance of the expression analysis. The set of conditions used for assessing the expression of *fhuA* and *smc01610* are given in table 5.4.

Table 5.4: Growth conditions for expression analysis of *fhuA* and *smc01610*

Conditions Required	Supplement Added	Supplement Concentration
Iron Replete	None	N/A
Iron Deplete	2,2'Dipyridyl	300 µM
Iron Replete, 1 µM Ferrichrome	Ferrichrome	1 µM
Iron Deplete, 1 µM Ferrichrome	2,2'Dipyridyl	300 µM
	Ferrichrome	1 µM
Iron Deplete, 100 nM Ferrichrome	2,2'Dipyridyl	300 µM
	Ferrichrome	100 nM
Iron Deplete, 10 nM Ferrichrome	2,2'Dipyridyl	300 µM
	Ferrichrome	10 nM

The ferrichrome concentration was varied as an anomalous result was obtained when gene expression was assessed in deplete conditions supplemented with 1 µM ferrichrome. This will be explored in more detail in the following section. In concurrence with the observed result for coprogen it was noticed that the 2,2'dipyridyl significantly reduced the growth rate. The growth rate was subsequently restored to levels close to that of the replete culture when supplemented with ferrichrome indicating the acquisition of iron through ferrichrome.

Cell growth in each condition was allowed to proceed to an OD_{600} reading of 0.6 was reached and at this point 15 ml of the cells was harvested and subjected to RNA extraction as previously described. The RNA samples were treated the same as the samples in the coprogen analysis. The cDNA reaction and subsequent qPCR analysis was carried out as previously described. As before the reference gene used was *gapA*

and *fhuP* was included as part of each analysis as an indicator to the internal iron status of the cell. All other methods of analysis and controls were implemented the same as in the coprogen system expression analysis.

5.5.1 Expression analysis of *fhuA* and *sma01610*

The involvement of iron limitation in the expression of *fhuA* and *smc01610* was determined by comparing the expression levels of a deplete culture to that of a replete culture. As carried out with the coprogen system analysis three genes of interest were monitored namely, *fhuA*, *sma01610* and *fhuP*. The results of this analysis are given in figure 5.6.

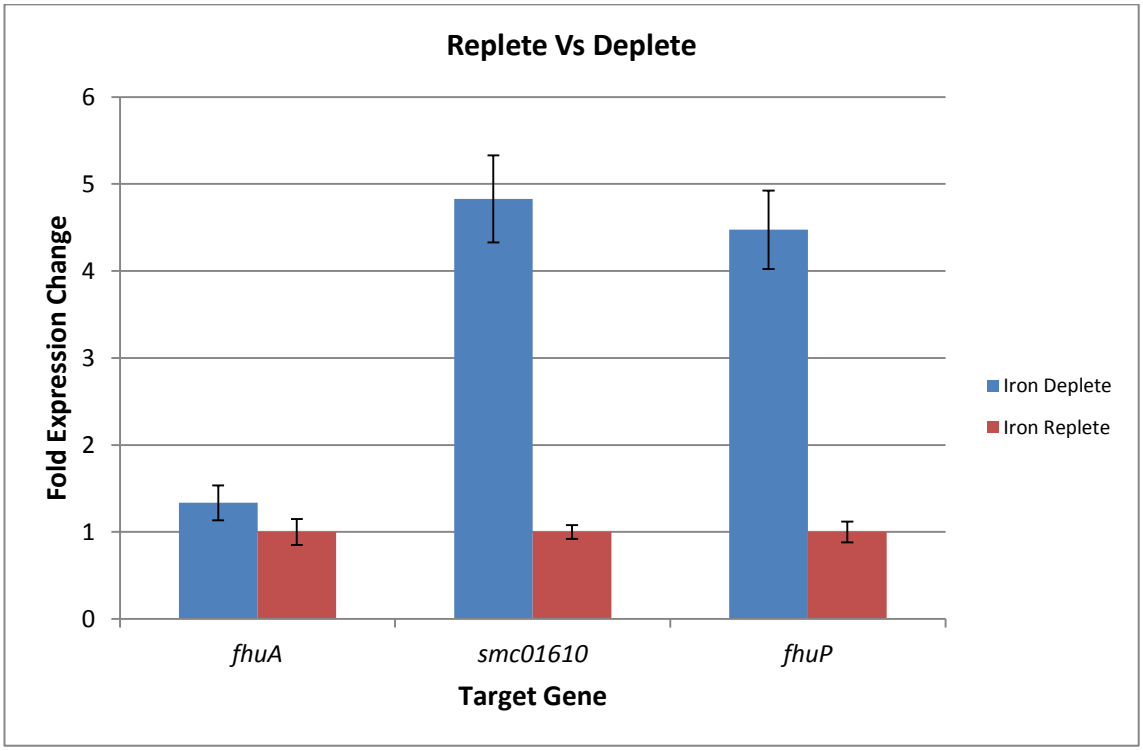


Figure 5.6: Expression analysis of *fhuA* and *smc01610* in iron deplete conditions

Expression of *fhuP* was up regulated in the iron deplete environment in comparison to an iron replete environment indicating that RirA was not repressing gene expression. There is a negligible change in expression of *fhuA* and there was a strong increase in the expression of *smc01610.* This analysis indicates that *fhuA* is not strongly influenced by iron limitation and that *smc01610* is influenced by iron limitation. The up regulation of *fhuP* is expected as it is known to be influenced by RirA and indicates that the culture was iron deprived.

As the iron deplete conditions did not show any significant effect on the expression of *fhuA* the effect exerted by ferrichrome on the expression of this gene set was determined. The iron replete condition expression levels was compared to the expression levels obtained from the iron replete conditions supplemented with 1 µM ferrichrome. The result of this analysis is shown in figure 5.7.

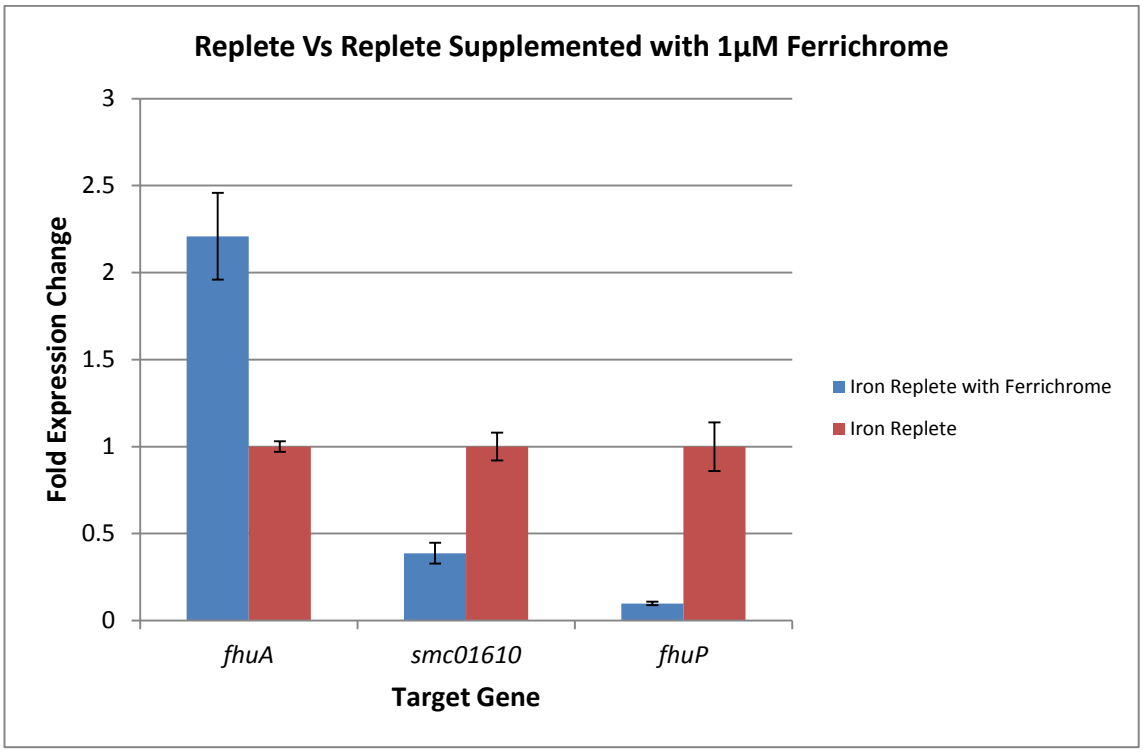


Figure 5.7: Expression analysis of *fhuA* and *smc01610* in iron replete conditions supplemented with 1 µM ferrichrome

As can be seen in this result set the expression of *fhuA* has increased in the presence of ferrichrome. The addition of ferrichrome results in the down regulation of *fhuP*, with a tenfold reduction compared to the levels observed in the replete culture which is not unexpected due to this also occurring in the coprogen system analysis. This is in keeping with the idea of ferrichrome providing an easily obtainable source of iron thus increasing the iron concentration in the cell therefore causing repression of *smc01610* and *fhuP*. The down regulation of *smc01610* linked to the observed up regulation in iron deplete conditions indicates that *smc01610* is influenced by cellular iron concentrations as this mimics the pattern of *fhuP*.

The up regulation of *fhuA* in the presence of ferrichrome is an indication that *fhuA* expression is directly linked to the presence of ferrichrome but does not show whether this is augmented under iron limitation. To investigate the role of iron in *fhuA* induction

the gene expression profile of this gene set was analysed in a deplete background supplemented with 1 µM ferrichrome. As before, the gene expression profile was normalised to iron replete conditions. The results of this analysis are given in figure 5.8.

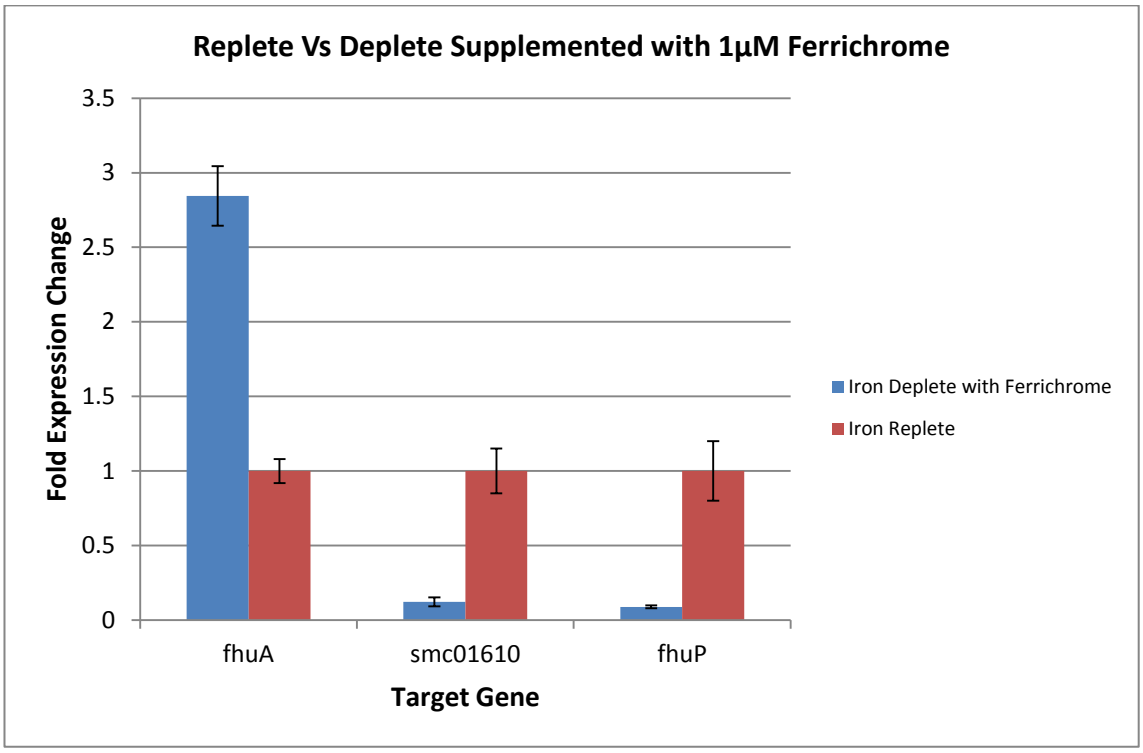


Figure 5.8: Expression analysis of *fhuA* and *smc01610* in iron deplete conditions supplemented with 1 µM ferrichrome

It is important to consider the expression of the *fhuP* control gene prior to discussing the effect of ferrichrome in an iron deplete environment on the expression of *fhuA*. The *fhuP* gene is down regulated to the same level as observed in the ferrichrome replete sample. This result taken by itself indicates that the culture was not iron deprived and appears to be fed iron in the form of ferrichrome. This is further indicated by comparing the gene set profile obtained in figure 5.7 to figure 5.8 which essentially follow the same pattern regardless of the presence of 2,2'dipyridyl. However, this result does provide more evidence for the expression of *fhuA* being directly influenced by ferrichrome in what appears to be independent of iron limitation. If RirA functioned to repress *fhuA* in an iron replete background then the levels of *fhuA* expression in figure 5.8 would follow the same pattern as *fhuP* but the opposite is observed.

As *fhuP* expression is down regulated in an iron deprived environment supplemented with 1 µM ferrichrome this indicates that the culture does not have to work to attain iron, which essentially indicates that ferrichrome is present far in excess of what is

required. To determine if ferrichrome overload is the cause of this anomaly this experiment was repeat with 100 nM ferrichrome and 10 nM ferrichrome. The results of the 100 nM ferrichrome concentration are given in figure 5.9.

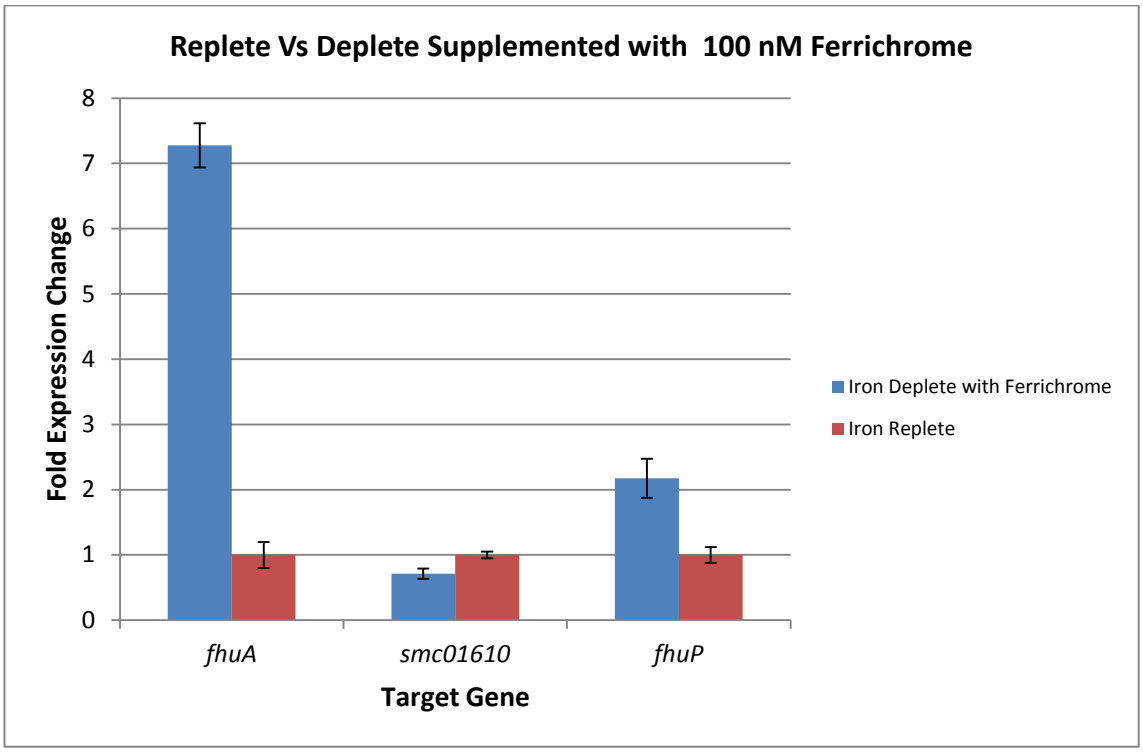


Figure 5.9: Expression analysis of fhuA and smc01610 in iron deplete conditions supplemented with 100 nM ferrichrome

The tenfold reduction in ferrichrome concentration has resulted in a drastic change of the expression profile. The induction of *fhuP* demonstrates that the culture is iron deprived and RirA is not fully repressing gene expression in agreement with that observed in figure 5.6. With the presence of the expected profile for *fhuP* expression the expression profiles of *fhuA* and *smc01610* can be correctly interpreted. The profile exhibited by *smc01610* suggests that it is not solely under the influence of iron concentration as if this was the case an up regulation would be observed. As *smc01610* is slightly down regulated this suggests a more complex role for this protein in the acquisition of ferrichrome. This role will be discussed and developed below but it indicates that Smc01610 may be regulating *fhuA* and *smc01610* expression differentially depending on the presence or absence of ferrichrome. As expected *fhuA* is up regulated and to a higher level than observed in the 1 µM ferrichrome deplete sample. This may suggest a complementary role for iron limitation in the expression of *fhuA* in an iron deplete environment but can also be explained through other

mechanisms. These possibilities will be discussed in full in the discussion section below.

As stated above, the profile of an iron deplete culture supplemented with 10 nM ferrichrome was also assessed to see if a similar expression pattern is attained. The results of this analysis are given in figure 5.10.

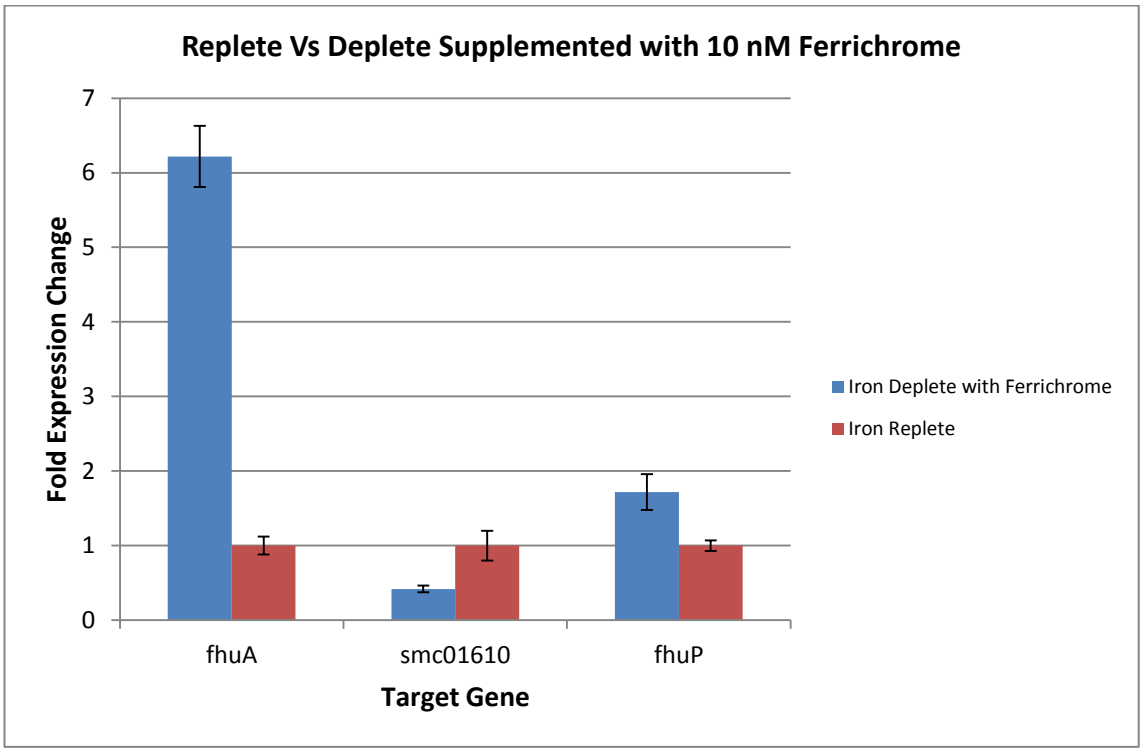


Figure 5.10: Expression analysis of *fhuA* and *smc01610* in iron deplete conditions supplemented with 10 nM ferrichrome

The results observed for the analysis of 10 nM ferrichrome in an iron deplete environment show the same pattern as an iron deplete environment with 100 nM ferrichrome present. It is interesting to note that *smc01610* is repressed once more while *fhuA* and *fhuP* are induced due to the inclusion of ferrichrome in an iron deprived environment. This may give further insights into the intricacies of Smc01610 activity. As *smc01610* was up regulated in iron delete conditions this indicated that it is influenced by the activity of RirA. However, this is the only incidence were it is seen to be induced, with all conditions where ferrichrome is present showing that it is down regulated. This strongly indicates that Smc01610 acts through RirA but has an additional level of auto-regulation. The above results indicate that in the absence of ferrichrome in iron deprivation it affects its own induction but when ferrichrome in present it switches preference and induces the expression of *fhuA*.

A possible explanation for the mode of action of Smc01610 will be presented in section 5.6. The following is a summary of the results discussed above.

5.5.2 Summary of the expression analysis of *fhuA* and *smc01610*

The analysis of the expression of *fhuA* and *smc01610* has yielded insights into how this two gene system could be regulating ferrichrome transport across the outer membrane and essentially into the cell. The *fhuP* gene again proved to be an important internal reference as it acted as a reporter of iron homeostasis in the cell. This internal control allowed for the identification of unexpected culture conditions such as the overabundance of ferrichrome in the 1 μM ferrichrome deplete culture which was of importance in this analysis.

The patterns observed for the expression of *fhuA* suggest that it is induced in the presence of ferrichrome with iron limitation playing an indirect role in its regulation. This conclusion is drawn from the lack of strong expression in an iron deplete culture. This conclusion is supported by data demonstrating the up regulation of *fhuA* in the presence of ferrichrome in iron replete conditions. If a lack of iron was influencing gene expression a profile similar to that observed in *fhuP* expression would be observed. Perhaps an argument for iron responsiveness for *fhuA* regulation could be made based on the results obtained in the 100 nM ferrichrome deplete and the 10 nM ferrichrome deplete cultures as *fhuA* is more strongly induced in these cultures. However, considering the role of each protein, whether putative like Smc01610 or known like FhuP another explanation becomes apparent.

Consider FhuP, it is known to be essential for ferrichrome acquisition as it sequesters it in the periplasm and delivers it to the inner membrane transporter complex HmuUV for internalisation (O Cuiv et al. 2008). Smc01610 is strongly indicated to be a member of the AraC-like regulator family which are known to interact with the substrates related to the genes they regulate i.e. ferrichrome and *fhuA* or *smc01610*. Bearing the role of these two proteins in mind an explanation for the increase in *fhuA* expression in the presence of lower levels of ferrichrome in an iron deplete environment becomes apparent. As Smc01610 requires ferrichrome to be in the cytoplasm to function it is dependent on the activity of FhuP as without FhuP ferrichrome cannot be internalised. In the 1 μM ferrichrome replete conditions the levels of FhuP are low due to sufficient intracellular iron which is also the case in the 1 μM ferrichrome deplete conditions. This would

result in nominal levels of ferrichrome gaining access to the cytoplasm as there would be very little FhuP in the periplasm to sequester it. This in turn would leave Smc01610 with low levels of ferrichrome substrate with which to interact. This would lead to a relatively weak induction of *fhuA* which is seen in figures 5.7 and figure 5.8. The profiles of the 100 nM ferrichrome deplete and 10 nM ferrichrome deplete conditions can be considered in this light. FhuP is induced in both of these samples which should result in greater levels of ferrichrome gaining access to the cytoplasm than in the 1 μM ferrichrome deplete sample even through ferrichrome is present at a concentration tenfold and hundredfold less in the external environment. This would result in more Smc01610-ferrichrome interaction which in turn should result in stronger expression of *fhuA*.

5.6 Possible modes of action for Sma1749 and Smc01610

To propose a definite mode of action for Sma1749 and Smc01610 would require characterisation of the binding activities of both of these proteins. This is beyond the scope of the work contained within this thesis. However, it is known that AraC-like regulators and indeed most transcriptional regulators require a DNA recognition sequence. This DNA sequence is generally arranged as a direct separated repeat of between 17 base pairs in length located proximal to the promoter region of the target gene (Schleif 2010; Wickstrum et al. 2007; Gallegos et al. 1997; Reeder and Schleif 1993). Analysis of the binding sites of the RhrA protein by EMSA revealed that a direct repeat of 6 bp separated by 15 bp constituted the binding site which suggests that variability exists in the required length the repeat sequence (Viguier 2005).

It is known that the original AraC displays differing affinity for binding sites depending on the presence or absence of arabinose with the I_1 and I_2 binding sites showing a consensus of TAGCxxxTxxxxCCxxA (Carra and Schleif 1993). This is known as the light switch mechanism as the protein switches location from a position that is unfavourable to expression to a position that induces gene expression. A simplified schematic of this process is presented in figure 5.11.

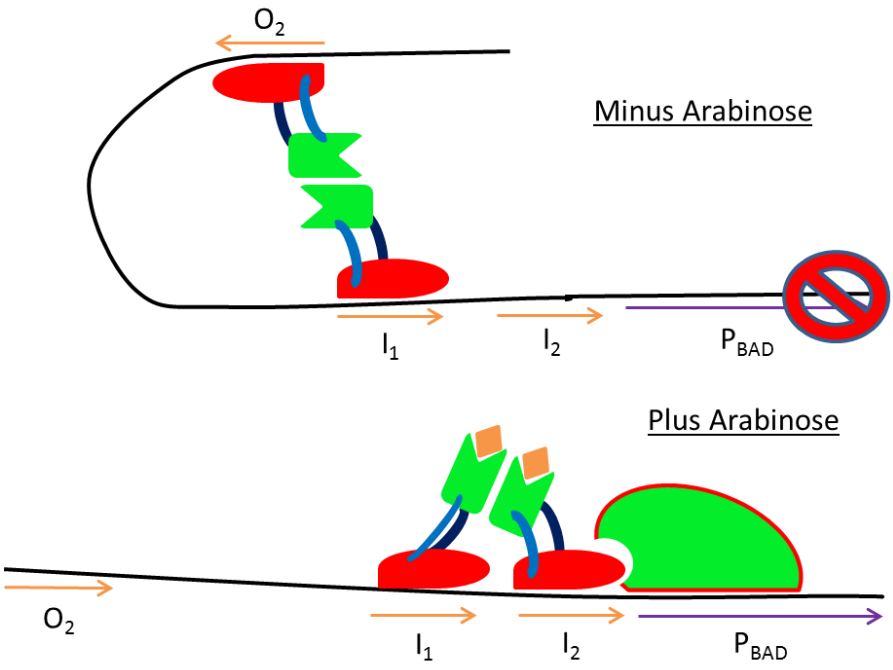


Figure 5.11: Schematic of the activity of AraC

The AraC protein is represented as a red DNA binding domain and a green substrate binding domain, arabinose is represented as an orange diamond and RNA polymerase is represented as green outlined in red. I_1, I_2 and O_2 are three different halfsites utilised by AraC and P_{BAD} is the promoter for the *araBAD* genes that encode for arabinose catabolism. In a minus arabinose situation AraC occupies the I_1 and O_2 half sites that are positioned ~200 bp apart. This introduces a DNA loop structure which hinders the recruitment of the RNA polymerase to the P_{BAD} which results in no expression of the *araBAD* genes. Once arabinose is detected by AraC it switches preference to the I_1 and I_2 binding sites which are located adjacent to one another as arabinose binding introduces a conformational change in the protein. The switch results in the DNA loop being opened which facilitates the recruitment of the RNA polymerase to the P_{BAD} promoter and once recruited the RNA polymerase can initiate transcription (Schleif 2010).

With the feature of a 6 or 7 bp direct repeat or larger consensus sequence for a half site recognition site for AraC-like proteins in mind, the regions immediately upstream of each of the four genes *sma1749, fhuE, smc01610* and *fhuA* were assessed for the presence of motifs resembling a 6-7 base pair separated repeat. Only two direct repeats located adjacent to the promoter will be considered as there may not be a DNA looping activity in Sma1749 and Smc01610.

5.6.1 Analysis of the promoter region of *fhuE*

Figure 5.12 is the DNA sequence defined by boundaries of the -200 position and +50 positions from the start codon of *fhuE*. This sequence extending to the stop codon of the gene was analysed for the possible presence of a promoter sequence by the Neural Network Promoter Prediction tool (Section 2.14).

TTCCGGCGAATCGTACAAAACCGGGGAAACATACAATACTAGACAAAACG
GCTCCAGTTAATAGCGGTTCGTGGGACAACGACTTCAAGCCGGGGTCGGT
TGAGCGAAGCTTTCATGCTTAGCGGACGGCCAGACGCGATGCAGTTCGGC
GACCTCAAAGGAGTTTGAGGCCCGCTTTCAGCACTTACGTCAATTGGAAC
ATGAACATGGACATCGATCAGACCGGCCCCTGGCCAACGAGATTACCGCT

Binding Site
Promoter Sequence
Transcriptional Start
Start Codon

Figure 5.12: Promter region of *fhuE*

There are two repeat sequences present in the region upstream of the predicted promoter sequence of *fhuE*. They comprise a 7 bp sequence and are positioned at position -146

and position -118 from the start of the predicted promoter. It is of interest that the transcriptional start site is marked as the start codon of the gene. This may not be correct as this would result in the absence of a ribosome binding site and may indicate that the actual promoter is closer to the putative binding sites.

5.6.2 Analysis of the promoter region of *sma1749*

Figure 5.13 is the DNA sequence defined by boundaries of the -200 position and +50 positions from the start codon of *sma1749*. This sequence extending to the stop codon of *sma1749* was analysed for the presence of a promoter sequence.

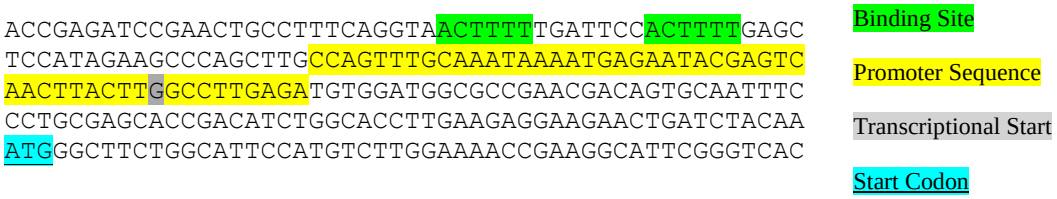


Figure 5.13: Promoter region of *sma1749*

There are two repeat sequences present in the region upstream of the predicted promoter sequence of *sma1749*. They comprise a 6 bp repeat and are positioned at positions -42 and position -29 from the start of the predicted promoter. The transcription start site is indicated at position -90 from the start codon which would allow for a ribosome binding site and other translational sites to be including the 5' untranslated region.

5.6.3 Analysis of the promoter region of *fhuA*

Figure 5.14 is the DNA sequence defined by boundaries of the -200 position and +50 positions from the start codon of *fhuA*. This sequence extending to the stop codon of *fhuA* was analysed for the presence of a promoter sequence.

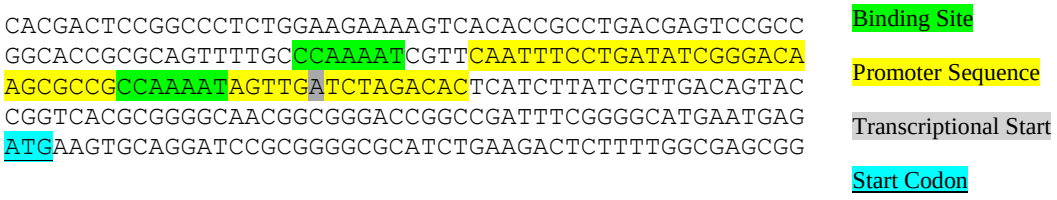


Figure 5.14: Promoter region of *fhuA*

There are two repeat sequences present in the region upstream of the predicted promoter sequence of *fhuA*. They comprise a 7 bp repeat sequence and are positioned at positions

–11 and +28 from the start of the predicted promoter sequence. The transcriptional start site is indicated at position -80 from the start codon which would allow for a ribosome binding site to be included upstream of the start codon.

5.6.4 Analysis of the promoter region of *smc01610*

Figure 5.15 is the DNA sequence defined by boundaries of the -200 position and +50 positions from the start codon of *smc01610*. This sequence extending to the stop codon of *smc01610* was analysed for the presence of a promoter sequence.

CCGGCCACCTTCTCCCCGCAGGCGGGGCGAAGGAGACTCGCGGCAGCCCT
CCGGCATTCCAAGGCGCGCGGCCGGCATGGCCGCGCGCCTTGGCATTGAA
GCGATTGAAAACCCGCGCGTTGCGCGCTAATAAACTTGACTTAAATGAGA
AGGATTGTTGCGGTTTTGCCTTGTAGCCGGCAGGACCGGAGAAAGCGACG
ATGACGTTCCAGCCGCGCATGCAGAACAGGATCAGCGGGTTTTCGATCAT

Binding Site

Promoter Sequence

Transcriptional Start

Start Codon

Figure 5.15: Promoter region of *smc01610*

The two repeat sequences are located at positions -51 and position -30 upstream of the predicted start of the promoter sequence and comprise a 7 bp repeat sequence. The transcriptional start site is indicated at position -59 upstream from the start codon which will allow for a ribosome binding site to be located upstream of the start codon.

5.6.5 Discussion of the putative binding sites and mode of action

The presence of putative DNA binding sites usually associated with AraC-like proteins near the promoter region of each of the genes indicates possible binding of Sma1749 to bind the region upstream of both *sma1749* and *fhuE* and for Smc01610 binding to the region upstream of both *smc01610* and *fhuA*. The predicted promoters for each of the genes are plausible with perhaps the predicted promoter for *fhuE* being positioned too close to the coding sequence. It must be noted that the promoter regions and transcription start sites annotated in the above sections are not definite. To absolutely define the transcriptional start and by extension the promoter, the length of the 5' untranslated region must be determined experimentally by a procedure such as primer extension or 5' rapid extension of cDNA ends (5' RACE). Alternatively this could be achieved through RNA sequencing if a more expansive analysis was preferred.

However, the identification of possible DNA binding sites presents the possibility of each of the promoter regions being a target of either Sma1749 or Smc01610. The change in sequence from that observed for *sma1749* to that of *fhuE* may represent a change of site affinities upon coprogen binding. The same could be occurring for the binding sites of *smc01610* and *fhuA.*

There is only slight conservation between the putative binding sites upstream of *sma1749* and *fhuE* with only the AC nucleotides being conserved between the halfsites. A similar situation is observed for the binding sites located upstream of *smc01610* and *fhuA* with a higher degree of conservation in the halfsite represented by CCNNNAT with NNN representing the variable region.

With the identification of putative binding sites, qPCR results and the known role of a number of the proteins involved it is possible to propose a model for how these two systems operate. In the absence of cellular protein levels and DNA binding assays this is only a best fit model as it cannot account for post-transcriptional regulation of protein levels and the activities of unknown relevant proteins such as the inner membrane transport system for coprogen. Also this model will be based on the assumption that if Sma1749 and Smc01610 are regulating their related systems, they are doing so as an activator not as a repressor. This assumption was made as the vast majority of homologues to AraC are activators not repressors (Gallegos et al. 1997).

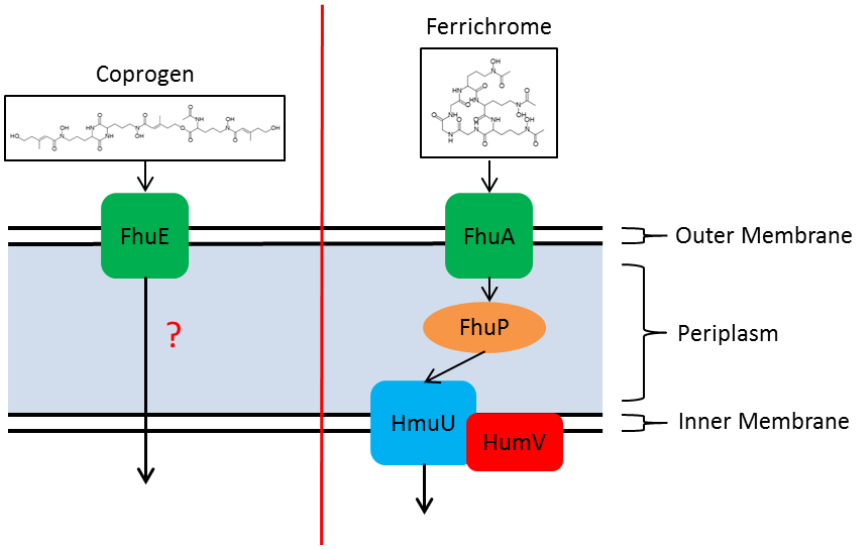


Figure 5.16: Schematic of the membrane transport of coprogen and ferrichrome

Figure 5.16 represents the known transport systems for the siderophore coprogen and ferrichrome. The outer membrane receptor for coprogen was identified as FhuE in chapter 4 of this thesis. The transport system for ferrichrome comprises the outer membrane receptor FhuA, the periplasmic binding protein FhuP and the inner membrane transport complex HmuUV (O Cuiv et al. 2008). These systems operate to internalise their cognate siderophores where they can be stripped of iron most likely through the activity of the siderophore reductase FhuF. Once internalised the siderophore interact with Sma1749 and Smc01610 for coprogen and ferrichrome respectively.

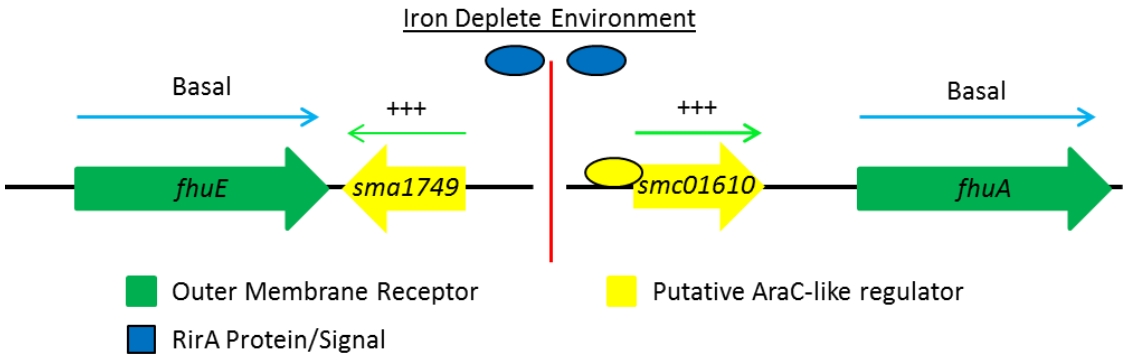


Figure 5.17: Proposed model for *fhuE* and *fhuA* regulation in iron deplete conditions

Figure 5.17 represents the proposed regulatory mechanism at both of the *fhuE* and *fhuA* loci when the cell is under iron stress with no cognate siderophore present. It is proposed that each of the AraC-like proteins Sma1749 and Smc01610 is positioned on the binding sites located upstream of their own promoter region. Sma1749 is proposed

to not be in a DNA binding conformation as coprogen is not present, therefore expression of *sma1749* is reliant on RirA de-repression. This will cause a build-up of Sma1749 and Smc01610 in the cell that can respond rapidly to the presence of the cognate siderophore. As the cells are iron deprived RirA is not in a DNA binding conformation thus is depicted as not being associated with DNA.

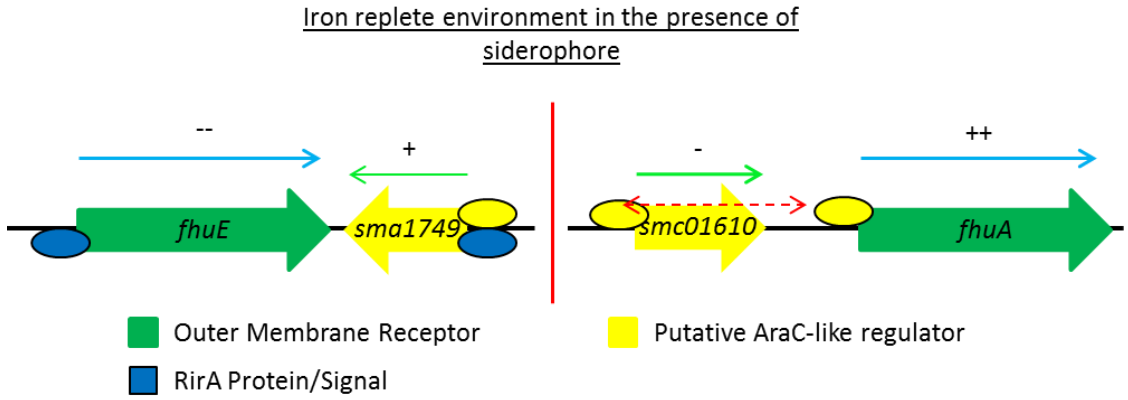


Figure 5.18: Proposed model for *fhuE* and *fhuA* regulation in iron replete conditions in the presence their cognate siderophore

With the addition of siderophore in an iron replete environment RirA is actively repressing its target genes of *sma1749*, *fhuE* and possibly *smc01610*. RirA is not depicted as binding to the promoter region of *smc01610* as it is uncertain whether the down regulation is due to the activity of RirA or due to a loss of affinity by Smc01610 for that binding site as discussed below. The activity of RirA is proposed to be countered by the activity of Sma1749 on its own promoter region which results in an increase in gene expression. RirA blocks the recruitment of Sma1749 to the promoter region of *fhuE* which prevents it from inducing gene expression even though coprogen is present. A different mechanism is proposed for the ferrichrome uptake system where the presence of ferrichrome results in a switch of preference of Smc01610 from its own promoter to the promoter of *fhuA*. As there is no activator of the *smc01610* gene expressing levels will drop as *fhuA* expression increases. This would result in an equilibrium forming between the levels of Smc01610 and ferrichrome in the cell. Smc01610 free of ferrichrome will induce *smc01610* and Smc01610 coordinated to ferrichrome will induce *fhuA*, this situation will result in high levels of intracellular ferrichrome causing a greater response towards the *fhuA* promoter and lower levels of ferrichrome resulting in induction of *smc01610*. This results in a fine balance in which

the strength of induction of either *smc01610* or *fhuA* depends on the levels of ferrichrome intracellularly with Smc01610 acting as a sensor.

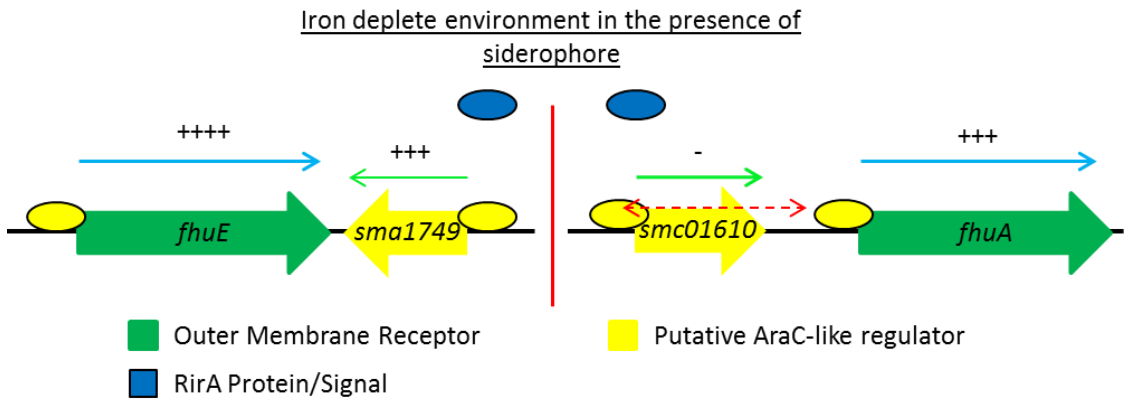


Figure 5.19: Proposed model for *fhuE* and *fhuA* regulation in iron deplete conditions in the presence their cognate siderophore

In iron deplete conditions in the presence each cognate siderophore expression of both of the outer membrane receptors *fhuE* and *fhuA* is strongly induced. This is due to the de-repression of RirA which removes it from the promoter of *fhuE*. This allows coprogen bound Sma1749 to bind the promoter and induce a strong response. This will allow more coprogen to enter the cell resulting in ample coprogen for Sma1749 to bind which will also result in greater expression of *sma1749*. This results in the large 20 fold induction of *fhuE* in these conditions along with the very strong induction in *sma1749* expression. This system would function as a form of positive feedback resulting in a large response to coprogen in the environment only if the cell is iron deprived.

As discussed previously for the ferrichrome system, a greater response was observed in *fhuA* in lower concentrations of ferrichrome. This is explained by the greater expression of *fhuP* which will result in increased levels of FhuP, and indeed HmuUV as they too are RirA regulated, allowing greater levels of ferrichrome to enter the cell. This will shift the previously mentioned equilibrium between Smc01610 and ferrichrome towards Smc01610 bound to ferrichrome which would result in greater expression of *fhuA* which is observed in the deplete conditions supplemented with 100 nM and 10 nM ferrichrome. This system would act as a sensitive sensor that induces the expression of *fhuA* regardless of the iron status of the cell. This would allow *S. meliloti* to immediately respond to the presence of ferrichrome with the level of response being

determined by the levels of the inner membrane transport system FhuP-HmuUV. As the inner membrane transport system is under the control of RirA this allows the uptake of ferrichrome to be controlled preventing a runaway induction of *fhuA* thus preventing over acquisition resulting in iron toxicity.

5.7 Conclusion

The qPCR analysis of the coprogen and ferrichrome acquisition systems presented here provides further insight into the intricacies of iron acquisition in *S. meliloti*. The coprogen system responds in a manner dependant on the cellular iron levels along with the presence of coprogen. The ferrichrome system responds to the presence of ferrichrome regardless of the cellular iron levels. However as the remaining components of the ferrichrome uptake system are under the control of RirA resulting in ferrichrome uptake still being influenced by iron deprivation.

The existence of such systems where the outer membrane receptor is only highly expressed in the presence of its cognate siderophore can afford many benefits to the host organism. Perhaps the most basic of these benefits is the conservation of energy and this occurs on two fronts. In a situation of iron deprivation any energy conservation mechanism is beneficial as many cellular processes require the presence of iron and reducing the number of genes needlessly induced would relieve a certain level of stress. Energy would also be conserved as a result of the strong up regulation of the outer membrane receptors upon encounter of the cognate siderophore as a high number of outer membrane receptors would greatly increase the chance of fulfilling the cells iron need directly from a xenosiderophore source. This would allow *S. meliloti* to relieve iron stress without having to produce a siderophore therefore eliminating the requirement to express a large range of genes. In consideration of the natural free living state of *S. meliloti* it is likely that xenosiderophores such as coprogen and ferrichrome could be satisfying iron acquisition to a large extent. This would give a competitive edge to *S. meliloti* as the organism with the most diverse means of acquiring iron would benefit from the vastly different siderophores present. Also, having a highly efficient mechanism for the uptake of xenosiderophores would allow *S. meliloti* to outcompete a neighbouring species for its own siderophore resulting in increased stress on the producing strain while simultaneously satisfying the iron requirement of *S. meliloti*. As mentioned in previous chapters ferrichrome and a number of its derivatives are present

at very high levels in numerous soil types, up to 10 nM in some instances (Essen et al. 2006; Prabhu, Biolchini and Boyer 1996). The qPCR results presented above showed that 10 nM ferrichrome is sufficient to induce expression of *fhuA* which indicates that ferrichrome could be a preferred means of acquiring iron in a free living state. Perhaps the induction of *fhuA* in iron replete conditions will allow *S. meliloti* to stockpile ferrichrome in the periplasmic space therefore preventing neighbouring organisms for acquiring it.

Aside from energy conservation another reason for only expressing an outer membrane receptor as required is to defend against phage attack and toxins. Perhaps the most prominent example of an outer membrane receptor facilitating phage attack is the first discovered namely FhuA of *E. coli*. Extensive analysis of FhuA has shown that it can be exploited by numerous substances such as phage T1, T5, φ80 and UC-1 and the toxins colicin M, microcin 25, albomycin and rifamycin CGP 4832. Indeed FhuA was originally named TonA as it was discovered to be the entry point of the T1 phage long before it was known to transport ferrichrome (Braun 2009). With the presence of outer membrane receptors leaving the cell open to attack by toxins and phage it would be an important survival adaptation to uncouple induction of the outer membrane receptor and the iron homoeostatic condition of the cell. This would position the *S. meliloti* where it could still benefit from the presence of the xenosiderophores coprogen and ferrichrome but also protect itself from phage attack in conditions of iron stress in the absence of these siderophores.

I would like to acknowledge the contribution of a final year project student, Edel Byrne for aiding with the setup of many of the reaction plates used as part of this study.

Concluding Remarks

6.1 Concluding Remarks

The degree of variability in iron acquisition methods is a reflection of the pivotal role it plays in bacterial survival. This is exemplified by the huge variation in siderophore structures, transport mechanisms and regulatory systems. The diversity of siderophore structure is discussed in detail in section 1.2, each with its own distinct properties with all but a few displaying remarkable affinities for ferric iron. As discussed in Chapter 3 numerous siderophores are acetylated with acetyl chains of various lengths that attribute these siderophores with amphipathic properties. This modification generated great interest for both the biosynthesis of these structures and also the biological significance of such appendages. Due to the hydrophobic nature of long acetyl chain groups these amphipathic siderophores have been shown to associate with biological membranes, partake in micelle formation and in some models have been shown to cross lipid bilayers by a flip-flop mechanism (Sandy and Butler 2009). The majority of the proposed functions of lipid structures on siderophores are centred on preventing siderophore diffusion in highly aqueous environments. Many models for the biosynthesis of these siderophores have been constructed however the key protein responsible for the addition of the lipid has yet to be identified. RhbG prior to characterisation represented one of the foremost candidates for the role of lipid addition based on its genome location and bioinformatic analysis. However, the results described herein showing intact rhizobactin 1021 being produced by an *rhbG* deletion mutant clearly demonstrates that the protein or proteins for lipid addition are encoded elsewhere on the genome.

As analysis of siderophore biosynthesis pathways has proved fruitless in identifying acetyltransferases involved in lipid addition it is prudent to look elsewhere for this function. Fatty acid biosynthesis has been studied extensively in *E. coli* and this pathway may act as a good model for fatty acid biosynthesis in *S. meliloti*. Two proteins that are of interest to rhizobactin 1021 biosynthesis that are present in the fatty acid biosynthesis pathway are AcpP and FabA. AcpP plays a central role in fatty acid biosynthesis as it is the acyl carrier protein around which the synthesis and elongation of fatty acid chains occurs. FabA plays the specific role of converting the saturated fatty acid decanol-ACP into the unsaturated fatty acids of decanyl-3-ACP or decanyl-2-ACP (Campbell and Cronan 2001). These proteins are of interest as mutagenesis of either protein should abolish production of the unsaturated fatty acid decanyl-2-CoA which is

the source of the lipid structure on rhizobactin 1021. Analysis of the *S. meliloti* genome for the presence of these genes identified two homologs putatively designated as *acpP* and *fabA*. These genes are arranged as members of two separate operons with the remaining genes in each operon encoding for fatty acid biosynthesis. To further the work contained within this thesis mutagenesis of both *acpP* and *fabA* would be the highest priority. If these proteins function as predicted this should abolish the production of rhizobactin 1021 but not necessarily siderophore production. If no siderophore is produced in a strain deficient in decanyl-2-CoA this would suggest that the core siderophore structure is assembled on the fatty acid. If an aberrant siderophore is formed where a different lipid is substituted for decanyl-2-CoA or if schizokinen is produced then this would suggest that the core structure is assembled first and the lipid or acetyl moiety added upon completion. Depending on the results of such an analysis various approaches could be undertaken to determine the protein interactions of these two proteins which may identify target proteins for the role of lipid addition. This protein analysis, along with bioinformatic analysis of other acetylated siderophore producing strains may allow the identification of a generic pathway or protein fulfilling siderophore acetylation in organisms producing amphipathic siderophores.

The identification of the outer membrane receptor for coprogen, FhuE further elaborates on the ability of *S. meliloti* to scavenge fungal siderophores for iron. This builds on the outer membrane receptors that have already been identified namely RhtA, ShmR, FhuA and FoxA encoded for the transport of rhizobactin 1021, haem, ferrichrome and ferrioxamine B respectively (Amarelle, O'Brian and Fabiano 2008; O Cuiv et al. 2008; Battistoni et al. 2002; Lynch et al. 2001). In addition to the discovery of FhuE the presence of a possible redundant inner membrane transport system outside of the system for haem, ferrichrome and ferrioxamine B demonstrates that there are further uncharacterised siderophore transport systems. The inner membrane transport systems currently identified in *S. meliloti* include the HmuTUV ABC transport complex for haem, FhuP-HmuUV ABC transport complex for both ferrichrome and ferrioxamine B, and RhtX for rhizobactin 1021 and structurally related siderophores (O Cuiv et al. 2008; O Cuiv et al. 2004). The putative ABC transporter complex encoded proximal to *fhuE*, *sma1746-sma1740* most likely encodes for coprogen transport and this along with the confirmation that there is a second coprogen inner membrane transport system indicates that there are at least two additional inner membrane transport systems for siderophore

uptake to those already characterised. The functionality of Sma1746-Sma1740 could be confirmed through complementation in a heterologous host such as *E. coli* or *P. aeruginosa*. However due to inherent differences between the organisms there is no certainty that the proteins would function correctly in a different organism. Observations made by O Cuiv et al. (2004) while investigating ferrioxamine B and other siderophore transport in *P. aeruginosa* showed that heterologous expression can be unreliable. Experiments where *foxA* from *Yersinia enterocolitica* and *foxB* from *P. aeruginosa* were present *in trans* resulted in the ability to utilise ferrioxamine B not being conferred to an *E. coli* host. Complementation of a *S. meliloti* strain deficient in ferrioxamine B transport across the inner membrane with *foxB* from *P. aeruginosa* did result in the restoration of ferrioxamine B utilisation. However, regardless of the possibility of the complementing proteins not functioning correctly, heterologous reconstitution of the possible coprogen transport system is the simplest method to investigate the function of Sma1746-Sma1740.

The discovery of coprogen and ferrichrome acting as effector molecules for the expression of *fhuE* and *fhuA* further demonstrates the complexity of the regulation of iron transport systems in *S. meliloti*. As discussed in summation of Chapter 5 there are many benefits to systems where outer membrane receptors are only highly expressed in the presence of their cognate siderophore. Such benefits include energy conservation both in the presence and absence of the cognate siderophore and protection against phage and toxins that rely on siderophore outer membrane receptors as an entry point. It is striking that in *S. meliloti* each of the outer membrane receptor genes *fhuA, fhuE, rhtA* and *shmR* are dependent on regulatory functions other than RirA suggesting that tight control of outer membrane expression is of paramount importance. Expression of *shmR* is controlled by a small regulatory protein known as HmuP which is transcribed as part of the operon encoding the inner membrane ABC transporter for haem arranged as *hmuPSTUV*. HmuS is thought to play a role in haemin degradation but this is unconfirmed. HmuP was shown to be vital for induction of *shmR* with knockout of *hmuP* resulting in no expression of *shmR* in iron deplete conditions with RirA functioning to repress expression in iron replete conditions (Amarelle et al. 2010; Viguier et al. 2005). This HmuP-ShmR system only responds to iron levels in the cell unlike the AraC-like regulation of *rhtA, fhuE* and *fhuA* which respond to the presence of

the cognate siderophore with or without addition regulation through RirA (Lynch et al. 2001).

Outer membrane receptors for siderophores have been identified as targets for "Trojan Horse" antibacterial methods. This method of controlling bacterial growth mimics naturally occurring siderophore-antibiotic compounds such as albomycin and ferrimycin. However, clinical implementation of siderophore-antibiotic derivatives has been slow as resistance readily occurs via mutations in the receptor or the TonB complex (Miethke and Marahiel 2007). AraC-like control of outer membrane receptors further increases the possibility of resistance to "Trojan Horse" antibiotics. Conjugation of an antibiotic to a siderophore may interfere with AraC-like binding to the siderophore resulting in a lack of induction of the outer membrane receptor. This would allow resistance to occur without the need for possibly deleterious changes to the receptor or TonB complex. To increase the viability of siderophore "Trojan Horse" antimicrobials such regulatory mechanisms must be considered when selecting the targeting siderophore.

References

Amarelle, V., Kozio, U., Rosconi, F., Noya, F., O'Brian, M.R. and Fabiano, E. 2010. A new small regulatory protein, HmuP, modulates haemin acquisition in *Sinorhizobium meliloti*. *Microbiology-Sgm*, 156, pp.1873-1882.

Amarelle, V., O'Brian, M.R. and Fabiano, E. 2008. ShmR is essential for utilization of heme as a nutritional iron source in *Sinorhizobium meliloti*. *Applied and Environmental Microbiology*, 74(20), pp.6473-6475.

Ampe, F., Kiss, E., Sabourdy, F. and Batut, J. 2003. Transcriptome analysis of *Sinorhizobium meliloti* during symbiosis. *Genome Biology*, 4(2), pp.R15.

Anderson, E.S., Paulley, J.T., Martinson, D.A., Gaines, J.M., Steele, K.H. and Roop,R.Martin,,II. 2011. The iron-responsive regulator Irr is required for wild-type expression of the gene encoding the heme transporter BhuA in *Brucella abortus* 2308. *Journal of Bacteriology*, 193(19), pp.5359-5364.

Andrews, S., Robinson, A. and Rodriguez-Quinones, F. 2003. Bacterial iron homeostasis. *FEMS Microbiology Reviews*, 27(2-3), pp.215-237.

Bagg, A. and Neilands, J. 1987. Ferric uptake regulation protein acts as a repressor, employing iron(ii) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochemistry*, 26(17), pp.5471-5477.

Barnett, M.J. and Fisher, R.F. 2006. Global gene expression in the rhizobial-legume symbiosis. *Symbiosis*, 42(1), pp.1-24.

Barnett, M., Fisher, R., Jones, T., Komp, C., Abola, A., Barloy-Hubler, F., Bowser, L., Capela, D., Galibert, F., Gouzy, J., Gurjal, M., Hong, A., Huizar, L., Hyman, R., Kahn, D., Kahn, M., Kalman, S., Keating, D., Palm, C., Peck, M., Surzycki, R., Wells, D., Yeh, K., Davis, R., Federspiel, N. and Long, S. 2001. Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *Proceedings of the National Academy of Sciences of the United States of America*, 98(17), pp.9883-9888.

Barnett, M., Tolman, C., Fisher, R. and Long, S. 2004. A dual-genome symbiosis chip for coordinate study of signal exchange and development in a prokaryote-host interaction. *Proceedings of the National Academy of Sciences of the United States of America,* 101(47), pp.16636-16641.

Barona-Gomez, F., Lautru, S., Francou, F., Leblond, P., Pernodet, J. and Challis, G.L. 2006. Multiple biosynthetic and uptake systems mediate siderophore-dependent iron acquisition in *Streptomyces coelicolor* A3(2) and *Streptomyces ambofaciens* ATCC 23877. *Microbiology-Sgm,* 152, pp.3355-3366.

Barry, S.M. and Challis, G.L. 2009. Recent advances in siderophore biosynthesis. *Current Opinion in Chemical Biology,* 13(2), pp.205-215.

Battistoni, F., Platero, R., Duran, R., Cervenansky, C., Battistoni, J., Arias, A. and Fabiano, E. 2002. Identification of an iron-regulated, hemin-binding outer membrane protein in *Sinorhizobium meliloti*. *Applied and Environmental Microbiology,* 68(12), pp.5877-5881.

Beasley, F.C., Vines, E.D., Grigg, J.C., Zheng, Q., Liu, S., Lajoie, G.A., Murphy, M.E.P. and Heinrichs, D.E. 2009. Characterization of staphyloferrin A biosynthetic and transport mutants in *Staphylococcus aureus*. *Molecular Microbiology,* 72(4), pp.947-963.

Benson, H., Boncompagni, E. and Guerinot, M. 2005. An iron uptake operon required for proper nodule development in the *Bradyrhizobium japonicum*-soybean symbiosis. *Molecular Plant-Microbe Interactions,* 18(9), pp.950-959.

Braun, V. 2009. FhuA (TonA), the career of a protein. *Journal of Bacteriology,* 191(11), pp.3431-3436.

Braun, V. and Hankte, K. 2013. The Tricky Ways Bacteria Cope with Iron Limitation, *IN:* Chakraborty, R., Braun, V., Hankte, K. and Cornelis. P. (Eds) *Iron Uptake in Bacteria with Emphasis on E. coli and Pseudomonas*. Vienna: Springer Briefs in Biometals, pp.31-66.

Braun, V., Hantke, K. and Koster, W. 1998. Bacterial iron transport: Mechanisms, genetics, and regulation. *Metal Ions in Biological Systems,* 35, pp.67-145.

Braun, V. and Mahren, S. 2005. Transmembrane transcriptional control (surface signalling) of the *Escherichia coli* Fec type. *FEMS Microbiology Reviews,* 29(4), pp.673-684.

Brewin, N. 2004. Plant cell wall remodelling in the rhizobium-legume symbiosis. *Critical Reviews in Plant Sciences,* 23(4), pp.293-316.

Brillet, K., Meksem, A., Lauber, E., Reimmann, C. and Cobessi, D. 2009. Use of an in-house approach to study the three-dimensional structures of various outer membrane proteins: Structure of the alcaligin outer membrane transporter FauA from *Bordetella pertussis*. *Acta Crystallographica Section D-Biological Crystallography,* 65, pp.326-331.

Buchanan, S., Smith, B., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., van der Helm, D. and Deisenhofer, J. 1999. Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*. *Nature Structural Biology,* 6(1), pp.56-63.

Budzikiewicz, H. 2010. Microbial Siderophores *IN:* Kinghorn, D. A., Falk, H. and Kobayashi, J. (Eds) *Progress in the Chemistry of Organic Natural Products.* Vienna: Springer-Verlay/Wien, pp.2-75.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J. and Wittwer, C.T. 2009. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry,* 55(4), pp.611-622.

Campbell, J. and Cronan, J. 2001. Bacterial fatty acid biosynthesis: Targets for antibacterial drug discovery. *Annual Review of Microbiology,* 55, pp.305-332.

Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dreano, S., Gloux, S., Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D., Pohl, T., Portetelle, D., Puhler, A., Purnelle, B., Ramsperger, U., Renard, C., Thebault, P., Vandenbol, M., Weidner, S. and Galibert, F. 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proceedings of the National Academy of Sciences of the United States of America,* 98(17), pp.9877-9882.

Capela, D., Filipe, C., Bobilk, C., Batut, J. and Bruand, C. 2006. *Sinorhizobium meliloti* differentiation during symbiosis with alfalfa: A transcriptomic dissection. *Molecular Plant-Microbe Interactions,* 19(4), pp.363-372.

Carter, R., Worsley, P., Sawers, G., Challis, G., Dilworth, M., Carson, K., Lawrence, J., Wexler, M., Johnston, A. and Yeoman, K. 2002. The *vbs* genes that direct synthesis of the siderophore vicibactin in *Rhizobium leguminosarum*: Their expression in other genera requires ECF sigma factor Rpol. *Molecular Microbiology,* 44(5), pp.1153-1166.

Carra, J. and Schleif, R. 1993. Variation of half-site organization and DNA looping by AraC protein. *Embo Journal,* 12(1), pp.35-44.

Challis, G. 2005. A widely distributed bacterial pathway for siderophore biosynthesis independent of nonribosomal peptide synthetases. *Chembiochem,* 6(4), pp.601-611.

Chao, T., Becker, A., Buhrmester, J., Puhler, A. and Weidner, S. 2004. The *Sinorhizobium meliloti fur* gene regulates, with dependence on mn(II), transcription of the *sitABCD* operon, encoding a metal-type transporter. *Journal of Bacteriology,* 186(11), pp.3609-3620.

Chao, T., Buhrmester, J., Hansmeier, N., Puhler, A. and Weidner, S. 2005. Role of the regulatory gene *rirA* in the transcriptional response of *Sinorhizobium meliloti* to iron limitation. *Applied and Environmental Microbiology,* 71(10), pp.5969-5982.

Chatfield, C.H., Mulhern, B.J., Viswanathan, V.K. and Cianciotto, N.P. 2012. The major facilitator superfamily-type protein LbtC promotes the utilization of the legiobactin siderophore by *Legionella pneumophila*. *Microbiology-Sgm,* 158, pp.721-735.

Chimento, D., Kadner, R. and Wiener, M. 2005. Comparative structural analysis of TonB-dependent outer membrane transporters: Implications for the transport cycle. *Proteins-Structure Function and Bioinformatics,* 59(2), pp.240-251.

Clarke, T., Braun, V., Winkelmann, G., Tari, L. and Vogel, H. 2002. X-ray crystallographic structures of the *Escherichia coli* periplasmic protein FhuD bound to hydroxamate-type siderophores and the antibiotic albomycin. *Journal of Biological Chemistry,* 277(16), pp.13966-13972.

Clarke, T., Ku, S., Dougan, D., Vogel, H. and Tari, L. 2000. The structure of the ferric siderophore binding protein FhuD complexed with gallichrome. *Nature Structural Biology,* 7(4), pp.287-291.

Cobessi, D., Celia, H., Folschweiller, N., Schalk, I., Abdallah, M. and Pattus, F. 2005. The crystal structure of the pyoverdine outer membrane receptor FpvA from *Pseudomonas aeruginosa* at 3.6 angstrom resolution. *Journal of Molecular Biology,* 347(1), pp.121-134.

Cobessi, D., Celia, H. and Pattus, F. 2005. Crystal structure at high resolution of ferric-pyochelin and its membrane receptor FptA from *Pseudomonas aeruginosa*. *Journal of Molecular Biology,* 352(4), pp.893-904.

Cobessi, D., Celia, H. and Pattus, F. 2004. Crystallization and X-ray diffraction analyses of the outer membrane pyochelin receptor FptA from pseudomonas aeruginosa. *Acta Crystallographica Section D-Biological Crystallography,* 60, pp.1919-1921.

Cobessi, D., Meksem, A. and Brillet, K. 2010. Structure of the heme/hemoglobin outer membrane receptor ShuA from *Shigella dysenteriae*: Heme binding by an induced fit mechanism. *Proteins-Structure Function and Bioinformatics,* 78(2), pp.286-294.

Crosa, J. and Walsh, C. 2002. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiology and Molecular Biology Reviews,* 66(2), pp.223-249.

Cress, W., Johnson, G. and Barton, L. 1986. The role of endomycorrhizal fungi in iron uptake by hilaria-jamesii. *Journal of Plant Nutrition,* 9(3-7), pp.547-556.

Dellagi, A., Brisset, M., Paulin, J. and Expert, D. 1998. Dual role of desferrioxamine in *Erwinia amylovora* pathogenicity. *Molecular Plant-Microbe Interactions,* 11(8), pp.734-742.

Denison, R.F. and Kiers, E.T. 2011. Life histories of symbiotic rhizobia and mycorrhizal fungi. *Current Biology,* 21(18), pp.775-785.

de Lorenzo, V., Bindereif, A., Paw, B. and Neilands, J. 1986. Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in *Escherichia coli* K-12. *Journal of Bacteriology,* 165(2), pp.570-578.

de Lorenzo, V. and NeilandS, J. 1986. Characterization of *iuca* and *iucc* genes of the aerobactin system of plasmid ColV-K30 in *Escherichia coli*. *Journal of Bacteriology,* 167(1), pp.350-355.

Devanathan, S. and Postle, K. 2007. Studies on colicin B translocation: FepA is gated by TonB. *Molecular Microbiology,* 65(2), pp.441-453.

Diaz-Mireles, E., Wexler, M., Todd, J., Bellini, D., Johnston, A. and Sawers, R. 2005. The manganese-responsive repressor Mur of *Rhizobium leguminosarum* is a member of the Fur-superfamily that recognizes an unusual operator sequence. *Microbiology-Sgm,* 151, pp.4071-4078.

Dilworth, M., Carson, K., Giles, R., Byrne, L. and Glenn, A. 1998. *Rhizobium leguminosarum* bv. viciae produces a novel cyclic trihydroxamate siderophore, vicibactin. *Microbiology-UK,* 144, pp.781-791.

Draper, R.C., Martin, L.W., Beare, P.A. and Lamont, I.L. 2011. Differential proteolysis of sigma regulators controls cell-surface signalling in *Pseudomonas aeruginosa*. *Molecular Microbiology,* 82(6), pp.1444-1453.

Emery, T. 1971. Role of ferrichrome as a ferric ionophore in *Ustilago sphaerogena*. *Biochemistry,* 10(8), pp.1483-1488.

Endriss, F., Braun, M., Killmann, H. and Braun, V. 2003. Mutant analysis of the *Escherichia coli* FhuA protein reveals sites of FhuA activity. *Journal of Bacteriology,* 185(16), pp.4683-4692.

Escamilla-Hernandez, R. and O'Brian, M.R. 2012. HmuP is a coactivator of irr-dependent expression of heme utilization genes in *Bradyrhizobium japonicum*. *Journal of Bacteriology,* 194(12), pp.3137-3143.

Escolar, L., Perez-Martin, J. and de Lorenzo, V. 1998. Binding of the Fur (ferric uptake regulator) repressor of *Escherichia coli* to arrays of the GATAAT sequence. *Journal of Molecular Biology,* 283(3), pp.537-547.

Essen, S., Bylund, D., Holmstrom, S., Moberg, M. and Lundstrom, U. 2006. Quantification of hydroxamate siderophores in soil solutions of podzolic soil profiles in Sweden. *Biometals,* 19(3), pp.269-282.

Fadeev, E., Luo, M. and Groves, J. 2004. Synthesis, structure, and molecular dynamics of gallium complexes of schizokinen and the amphiphilic siderophore acinetoferrin. *Journal of the American Chemical Society,* 126(38), pp.12065-12075.

Feistner, G., Stahl, D. and Gabrik, A. 1993. Proferrioxamine siderophores of *Erwinia amylovora* - a capillary liquid-chromatographic electrospray tandem mass-spectrometric study. *Organic Mass Spectrometry,* 28(3), pp.163-175.

Ferguson, A., Braun, V., Fiedler, H., Coulton, J., Diederichs, K. and Welte, W. 2000. Crystal structure of the antibiotic albomycin in complex with the outer membrane transporter FhuA. *Protein Science,* 9(5), pp.956-963.

Ferguson, A., Chakraborty, R., Smith, B., Esser, L., van der Helm, D. and Deisenhofer, J. 2002. Structural basis of gating by the outer membrane transporter FecA. *Science,* 295(5560), pp.1715-1719.

Ferguson, A., Hofmann, E., Coulton, J., Diederichs, K. and Welte, W. 1998. Siderophore-mediated iron transport: Crystal structure of FhuA with bound lipopolysaccharide. *Science,* 282(5397), pp.2215-2220.

Fetherston, J.D., Kirillina, O., Bobrov, A.G., Paulley, J.T. and Perry, R.D. 2010. The yersiniabactin transport system is critical for the pathogenesis of bubonic and pneumonic plague. *Infection and Immunity*, 78(5), pp.2045-2052.

Fiedler, H., Krastel, P., Muller, J., Gebhardt, K. and Zeeck, A. 2001. Enterobactin: The characteristic catecholate siderophore of enterobacteriaceae is produced by *Streptomyces species*. *FEMS Microbiology Letters*, 196(2), pp.147-151.

Fillat, M.F. 2014. The FUR (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators. *Archives of Biochemistry and Biophysics*, 546, pp.41-52.

Finan, T.M., Kunkel, B., Devos, G.F. and Signer, E.R. 1986. 2nd Symbiotic Megaplasmid in *Rhizobium meliloti* Carrying Exopolysaccharide and Thiamine Synthesis Genes. *Journal of Bacteriology*, 167(1), pp.66-72.

Finan, T., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorholter, F., Hernandez-Lucas, I., Becker, A., Cowie, A., Gouzy, J., Golding, B. and Puhler, A. 2001. The complete sequence of the 1,683-kb pSymB megaplasmid from the N-2-fixing endosymbiont *Sinorhizobium meliloti*. *Proceedings of the National Academy of Sciences of the United States of America*, 98(17), pp.9889-9894.

Fischer, H. 1994. Genetic-regulation of nitrogen-fixation in rhizobia. *Microbiological Reviews*, 58(3), pp.352-386.

Flaten, T.P., Aaseth, J., Andersen, O. and Kontoghiorghes, G.J. 2012. Iron mobilization using chelation and phlebotomy. *Journal of Trace Elements in Medicine and Biology*, 26(2-3), pp.127-130.

Funahashi, T., Tanabe, T., Maki, J., Miyamoto, K., Tsujibo, H. and Yamamoto, S. 2013. Identification and characterization of a cluster of genes involved in biosynthesis and transport of acinetoferrin, a siderophore produced by *Acinetobacter haemolyticus* ATCC 17906(T). *Microbiology-Sgm*, 159, pp.678-690.

Gage, D. 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiology and Molecular Biology Reviews*, 68(2), pp.280-300.

Gage, D. 2002. Analysis of infection thread development using gfp- and DsRed-expressing *Sinorhizobium meliloti*. *Journal of Bacteriology*, 184(24), pp.7042-7046.

Galibert, F., Finan, T., Long, S., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R., Dreano, S., Federspiel, N., Fisher, R., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R., Jones, T., Kahn, D., Kahn, M., Kalman, S., Keating, D., Kiss, E., Komp, C., Lalaure, V., Masuy, D., Palm, C., Peck, M., Pohl, T., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thebault, P., Vandenbol, M., Vorholter, F., Weidner, S., Wells, D., Wong, K., Yeh, K. and Batut, J. 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science*, 293(5530), pp.668-672.

Gallegos, M., Schleif, R., Bairoch, A., Hofmann, K. and Ramos, J. 1997. AraC/XylS family of transcriptional regulators. *Microbiology and Molecular Biology Reviews*, 61(4), pp.393-410.

Geurts, R., Fedorova, E. and Bisseling, T. 2005. Nod factor signaling genes and their function in the early stages of rhizobium infection. *Current Opinion in Plant Biology*, 8(4), pp.346-352.

Gibson F. and Magrath, D. I. 1969. The isolation and characterization of a hydroxamic acid (aerobactin) formed by aerobacter aerogenes 62-I. *Biochimica Et Biophysica Acta-General Subjects*, 192(2), pp.175-184.

Glazebrook, J., Ichige, A. and Walker, G. 1993. A *Rhizobium meliloti* homolog of the *Escherichia coli* peptide antibiotic transport protein Sbma is essential for bacteroid development. *Genes & Development*, 7(8), pp.1485-1497.

Gledhill, M. 2001. Electrospray ionisation-mass spectrometry of hydroxamate siderophores. *Analyst*, 126(8), pp.1359-1362.

Greenwald, J., Nader, M., Celia, H., Gruffaz, C., Geoffroy, V., Meyer, J., Schalk, I.J. and Pattus, F. 2009. FpvA bound to non-cognate pyoverdines: Molecular basis of siderophore recognition by an iron transporter. *Molecular Microbiology*, 72(5), pp.1246-1259.

Gresock, M.G., Savenkova, M.I., Larsen, R.A., Ollis, A.A. and Postle, K. 2011. Death of the TonB shuttle hypothesis. *Frontiers in Microbiology*, 2, pp.206.

Gutjahr, C. and Parniske, M. 2013. Cell and developmental biology of arbuscular mycorrhiza symbiosis. *Annual Review of Cell and Developmental Biology*, 29, pp.593-617.

Hamza, I., Chauhan, S., Hassett, R. and O'Brian, M. 1998. The bacterial Irr protein is required for coordination of heme biosynthesis with iron availability. *Journal of Biological Chemistry*, 273(34), pp.21669-21674.

Hamza, I., Qi, Z., King, N. and O'Brian, M. 2000. Fur-independent regulation of iron metabolism by Irr in *Bradyrhizobium japonicum*. *Microbiology-Uk*, 146, pp.669-676.

Hantke, K. 2002. Members of the Fur protein family regulate iron and zinc transport in *E. coli* and characteristics of the fur-regulated FhuF protein. *Journal of Molecular Microbiology and Biotechnology*, 4(3), pp.217-222.

Hantke, K. 1983. Identification of an iron uptake system specific for coprogen and rhodotorulic acid in *Escherichia coli*-K12. *Molecular & General Genetics*, 191(2), pp.301-306.

Harris, W., Carrano, C. and Raymond, K. 1979. Coordination chemistry of microbial iron transport compounds .16. isolation, characterization, and formation-constants of ferric aerobactin. *Journal of the American Chemical Society*, 101(10), pp.2722-2727.

Heinrichs, D. and Poole, K. 1996. PchR, a regulator of ferripyochelin receptor gene (*fptA*) expression in *Pseudomonas aeruginosa*, functions both as an activator and as a repressor. *Journal of Bacteriology*, 178(9), pp.2586-2592.

Hibbing, M.E. and Fuqua, C. 2011. Antiparallel and interlinked control of cellular iron levels by the Irr and RirA regulators of *Agrobacterium tumefaciens*. *Journal of Bacteriology*, 193(14), pp.3461-3472.

Hider, R.C. and Kong, X. 2010. Chemistry and biology of siderophores. *Natural Product Reports*, 27(5), pp.637-657.

Koebnik, R., Locher, K. and Van Gelder, P. 2000. Structure and function of bacterial outer membrane proteins: Barrels in a nutshell. *Molecular Microbiology,* 37(2), pp.239-253.

Hohle, T.H. and O'Brian, M.R. 2010. Transcriptional control of the *bradyrhizobium japonicum irr* gene requires repression by Fur and antirepression by Irr. *Journal of Biological Chemistry,* 285(34), pp.26074-26080.

Hohle, T.H., Franck, W.L., Stacey, G. and O'Brian, M.R. 2011. Bacterial outer membrane channel for divalent metal ion acquisition. *Proceedings of the National Academy of Sciences of the United States of America,* 108(37), pp.15390-15395.

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop II, R. M. and Peterson K. M. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene,* 166, pp.175-176.

Hunter, S., Jones, P., Mitchell, A., Apweiler, R., Attwood, T.K., Bateman, A., Bernard, T., Binns, D., Bork, P., Burge, S., de Castro, E., Coggill, P., Corbett, M., Das, U., Daugherty, L., Duquenne, L., Finn, R.D., Fraser, M., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., McAnulla, C., McDowall, J., McMenamin, C., Mi, H., Mutowo-Muellenet, P., Mulder, N., Natale, D., Orengo, C., Pesseat, S., Punta, M., Quinn, A.F., Rivoire, C., Sangrador-Vegas, A., Selengut, J.D., Sigrist, C.J.A., Scheremetjew, M., Tate, J., Thimmajanarthanan, M., Thomas, P.D., Wu, C.H., Yeats, C. and Yong, S. 2012. InterPro in 2011: New developments in the family and domain prediction database. *Nucleic Acids Research,* 40(1), pp.306-312.

Ikai, H. and Yamamoto, S. 1997. Identification and analysis of a gene encoding L-2,4-diaminobutyrate:2-ketoglutarate 4-aminotransferase involved in the 1,3-diaminopropane production pathway in *Acinetobacter baumannii*. *Journal of Bacteriology,* 179(16), pp.5118-5125.

Ito, Y. and Butler, A. 2005. Structure of synechobactins, new siderophores of the marine cyanobacterium *Synechococcus sp.* PCC 7002. *American Society of Limnology and Oceanography, Inc,* 50(6), pp.1918-1923.

Jacobs, C., Huang, L., Bartowsky, E., Normark, S. and Park, J. 1994. Bacterial-cell wall recycling provides cytosolic muropeptides as effectors for beta-lactamase induction. *Embo Journal,* 13(19), pp.4684-4694.

Johnston, A.W.B., Todd, J.D., Curson, A.R., Lei, S., Nikolaidou-Katsaridou, N., Gelfand, M.S. and Rodionov, D.A. 2007. Living without Fur: The subtlety and complexity of iron-responsive gene regulation in the symbiotic bacterium rhizobium and other alpha-proteobacteria. *Biometals,* 20(3-4), pp.501-511.

Kachadourian, R., Dellagi, A., Laurent, J., Bricard, L., Kunesch, G. and Expert, D. 1996. Desferrioxamine-dependent iron transport in *Erwinia amylovora* CFBP1430: Cloning of the gene encoding the ferrioxamine receptor FoxR. *Biometals,* 9(2), pp.143-150.

Keen, N.T., Tamaki, S., Kobayashi, D. and Trollinger, D. 1988. Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene,* 70(1), pp.191-197.

Keogh, D. 2008. Haem and Xenosiderophore Mediated Iron Acquisition by *Sinorhizobium meliloti* and *Pseudomonas aeruginosa.* PhD Thesis. Dublin City University.

Koebnik, R., Locher, K. and Van Gelder, P. 2000. Structure and function of bacterial outer membrane proteins: Barrels in a nutshell. *Molecular Microbiology,* 37(2), pp.239-253.

Konetschnyrapp, S., Jung, G., Meiwes, J. and Zahner, H. 1990. Staphyloferrin-a - a structurally new siderophore from *Staphylococci*. *European Journal of Biochemistry,* 191(1), pp.65-74.

Krewulak, K.D. and Vogel, H.J. 2011. TonB or not TonB: Is that the question? *Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire,* 89(2), pp.87-97.

Krewulak, K.D. and Vogel, H.J. 2008. Structural biology of bacterial iron uptake. *Biochimica Et Biophysica Acta-Biomembranes,* 1778(9), pp.1781-1804.

Lee, J. and Helmann, J.D. 2007. Functional specialization within the Fur family of metalloregulators. *Biometals,* 20(3-4), pp.485-499.

Leong, S. and Winkelmann, G. 1998. Molecular biology of iron transport in fungi. *Metal Ions in Biological Systems,* 35, pp.147-186.

Letoffe, S., Delepelaire, P. and Wandersman, C. 2006. The housekeeping dipeptide permease is the *Escherichia coli* heme transporter and functions with two optional peptide binding proteins. *Proceedings of the National Academy of Sciences of the United States of America,* 103(34), pp.12891-12896.

Livak, K. and Schmittgen, T. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-delta delta C) method. *Methods,* 25(4), pp.402-408.

Locher, K., Rees, B., Koebnik, R., Mitschler, A., Moulinier, L., Rosenbusch, J. and Moras, D. 1998. Transmembrane signaling across the ligand-gated FhuA receptor: Crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell,* 95(6), pp.771-778.

Luo, M., Fadeev, E. and Groves, J. 2005. Membrane dynamics of the amphiphilic siderophore, acinetoferrin. *Journal of the American Chemical Society,* 127(6), pp.1726-1736.

Lynch, D., O'Brien, J., Welch, T., Clarke, P., O Cuiv, P, Crosa, J.H. and O'Connell, M. 2001. Genetic organization of the region encoding regulation, biosynthesis, and transport of rhizobactin 1021, a siderophore produced by *Sinorhizobium meliloti*. *Journal of Bacteriology,* 183(8), pp.2576-2585.

Ma, L., Kaserer, W., Annamalai, R., Scott, D.C., Jin, B., Jiang, X., Xiao, Q., Maymani, H., Massis, L.M., Ferreira, L.C.S., Newton, S.M.C. and Klebba, P.E. 2007. Evidence of ball-and-chain transport of ferric enterobactin through FepA. *Journal of Biological Chemistry,* 282(1), pp.397-406.

Mademidis, A., Killmann, H., Kraas, W., Flechsler, I., Jung, G. and Braun, V. 1997. ATP-dependent ferric hydroxamate transport system in *Escherichia coli*: Periplasmic FhuD interacts with a periplasmic and with a transmembrane cytoplasmic region of the integral membrane protein FhuB, as revealed by competitive peptide mapping. *Molecular Microbiology,* 26(5), pp.1109-1123.

Mademidis, A. and Koster, W. 1998. Transport activity of FhuA, FhuC, FhuD, and FhuB derivatives in a system free of polar effects, and stoichiometry of components

involved in ferrichrome uptake. *Molecular and General Genetics,* 258(1-2), pp.156-165.

Martin, J., Ito, Y., Homann, V., Haygood, M. and Butler, A. 2006. Structure and membrane affinity of new amphiphilic siderophores produced by *Ochrobactrum sp* SP18. *Journal of Biological Inorganic Chemistry,* 11(5), pp.633-641.

Martinez, J., Carter-Franklin, J., Mann, E., Martin, J., Haygood, M. and Butler, A. 2003. Structure and membrane affinity of a suite of amphiphilic siderophores produced by a marine bacterium. *Proceedings of the National Academy of Sciences of the United States of America,* 100(7), pp.3754-3759.

Martinez, J., Zhang, G., Holt, P., Jung, H., Carrano, C., Haygood, M. and Butler, A. 2000. Self-assembling amphiphilic siderophores from marine bacteria. *Science,* 287(5456), pp.1245-1247.

Masse, E. and Gottesman, S. 2002. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America,* 99(7), pp.4620-4625.

McHugh, J., Rodriguez-Quinones, F., Abdul-Tehrani, H., Svistunenko, D., Poole, R., Cooper, C. and Andrews, S. 2003a. Global iron-dependent gene regulation in *Escherichia coli* - A new mechanism for iron homeostasis. *Journal of Biological Chemistry,* 278(32), pp.29478-29486.

Meade, H.M., Long, S.R., Ruvkun, G.B., Brown, S.E. and Ausubel, F.M. 1982. Physical and genetic-characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *Journal of Bacteriology,* 149(1), pp.114-122.

Mei, B., Budde, A. and Leong, S. 1993. *sid1*, a gene initiating siderophore biosynthesis in *Ustilago maydis* - molecular characterization, regulation by iron, and role in phytopathogenicity. *Proceedings of the National Academy of Sciences of the United States of America,* 90(3), pp.903-907.

Mey, A., Wyckoff, E., Oglesby, A., Rab, E., Taylor, R. and Payne, S. 2002. Identification of the *Vibrio cholerae* enterobactin receptors VctA and IrgA: IrgA is not required for virulence. *Infection and Immunity,* 70(7), pp.3419-3426.

Michel, L., Bachelard, A. and Reimmann, C. 2007. Ferripyochelin uptake genes are involved in pyochelin-mediated signalling in *Pseudomonas aeruginosa*. *Microbiology-Sgm,* 153, pp.1508-1518.

Miethke, M. and Marahiel, M.A. 2007. Siderophore-based iron acquisition and pathogen control. *Microbiology and Molecular Biology Reviews,* 71(3), pp.413-451.

Miethke, M., Pierik, A.J., Peuckert, F., Seubert, A. and Marahiel, M.A. 2011. Identification and characterization of a novel-type ferric siderophore reductase from a gram-positive extremophile. *Journal of Biological Chemistry,* 286(3), pp.2245-2260.

Muller, G. and Raymond, K. 1984. Specificity and mechanism of ferrioxamine-mediated iron transport in *Streptomyces pilosus*. *Journal of Bacteriology,* 160(1), pp.304-312.

Mullis, K., Pollack, J. and Neilands, J. 1971. Structure of schizokinen, an iron-transport compound from *Bacillus megaterium*. *Biochemistry,* 10(26), pp.4894-4898.

Murray, J.D. 2011. Invasion by invitation: Rhizobial infection in legumes. *Molecular Plant-Microbe Interactions,* 24(6), pp.631-639.

Naka, H., Liu, M. and Crosa, J.H. 2013. Two ABC transporter systems participate in siderophore transport in the marine pathogen *Vibrio anguillarum* 775 (pJM1). *FEMS Microbiology Letters,* 341(2), pp.79-86.

Neilands, J. 1981. Microbial iron compounds. *Annual Review of Biochemistry,* 50, pp.715-731.

Nienaber, A., Hennecke, H. and Fischer, H. 2001. Discovery of a haem uptake system in the soil bacterium *Bradyrhizobium japonicum*. *Molecular Microbiology,* 41(4), pp.787-800.

Nikaido, H. 2003. Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews,* 67(4), pp.593-656.

O Cuiv, P, Clarke, P., Lynch, D. and O'Connell, M. 2004. Identification of *rhtX* and *fptX*, novel genes encoding proteins that show homology and function in the utilization of the siderophores rhizobactin 1021 by *Sinorhizobium meliloti* and pyochelin by

Pseudomonas aeruginosa, respectively. *Journal of Bacteriology,* 186(10), pp.2996-3005.

O Cuiv, P, Keogh, D., Clarke, P. and O'Connell, M. 2008. The *hmuUV* genes of *Sinorhizobium meliloti* 2011 encode the permease and ATPase components of an ABC transport system for the utilization of both haem and the hydroxamate siderophores, ferrichrome and ferrioxamine B. *Molecular Microbiology,* 70(5), pp.1261-1273.

O Cuiv, P, Keogh, D., Clarke, P. and O'Connell, M. 2007. FoxB of *Pseudomonas aeruginosa* functions in the utilization of the xenosiderophores ferrichrome, ferrioxamine B, and schizokinen: Evidence for transport redundancy at the inner membrane. *Journal of Bacteriology,* 189(1), pp.284-287.

O'Brian, R. M. and Fabiano, E. 2010. Mechanisms and Regulation of Iron Homeostasis in the Rhizobia *IN:* Cornelis, P. and Andrews, C. S. (Eds) *Iron Uptake and Homoeostasis in Microorganisms*. Norfolk: Caister Academic Press, pp.37-63.

O'Brien, I.G., Cox, G.B. and Gibson, F. 1970. Biologically active compounds containing 2,3-dihydroxybenzoic acid and serine formed by *Escherichia coli*. *Biochim Biophys Acta,* 201(3), pp.453-60.

Okujo, N., Sakakibara, Y., Yoshida, T. and Yamamoto, S. 1994. Structure of acinetoferrin, a new citrate-based dihydroxamate siderophore from *Acinetobacter haemolyticus*. *Biometals,* 7(2), pp.170-176.

Oldroyd, G. and Downie, J. 2004. Calcium, kinases and nodulation signalling in legumes. *Nature Reviews Molecular Cell Biology,* 5(7), pp.566-576.

Ong, S., Peterson, T. and Neilands, J. 1979. Agrobactin, a siderophore from *Agrobacterium tumefaciens*. *Journal of Biological Chemistry,* 254(6), pp.1860-1865.

Ott, T., van Dongen, J., Gunther, C., Krusell, L., Desbrosses, G., Vigeolas, H., Bock, V., Czechowski, T., Geigenberger, P. and Udvardi, M. 2005. Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. *Current Biology,* 15(6), pp.531-535.

Patel, P., Song, L. and Challis, G.L. 2010. Distinct extracytoplasmic siderophore binding proteins recognize ferrioxamines and ferricoelichelin in *Streptomyces coelicolor* A3(2). *Biochemistry,* 49(37), pp.8033-8042.

Perret, X., Staehelin, C. and Broughton, W. 2000. Molecular basis of symbiotic promiscuity. *Microbiology and Molecular Biology Reviews,* 64(1), pp.180-201.

Perry, R., Balbo, P., Jones, H., Fetherston, J. and DeMoll, E. 1999. Yersiniabactin from *Yersinia pestis*: Biochemical characterization of the siderophore and its role in iron transport and regulation. *Microbiology-Uk,* 145, pp.1181-1190.

Persmark, M., Pittman, P., Buyer, J., Schwyn, B., Gill, P. and Neilands, J. 1993. Isolation and structure of rhizobactin-1021, a siderophore from the alfalfa symbiont *Rhizobium meliloti* 1021. *Journal of the American Chemical Society,* 115(10), pp.3950-3956.

Pfeifer, B., Wang, C., Walsh, C. and Khosla, C. 2003. Biosynthesis of yersiniabactin, a complex polyketide-nonribosomal peptide, using *Escherichia coli* as a heterologous host. *Applied and Environmental Microbiology,* 69(11), pp.6698-6702.

Platero, R., Peixoto, L., O'Brian, M. and Fabiano, E. 2004. Fur is involved in manganese-dependent regulation of *mntA* (*sitA*) expression in *Sinorhizobium meliloti*. *Applied and Environmental Microbiology,* 70(7), pp.4349-4355.

Plessner, O., Klapatch, T. and Guerinot, M. 1993. Siderophore utilization by bradyrhizobium-japonicum. *Applied and Environmental Microbiology,* 59(5), pp.1688-1690.

Pohl, E., Haller, J., Mijovilovich, A., Meyer-Klaucke, W., Garman, E. and Vasil, M. 2003. Architecture of a protein central to iron homeostasis: Crystal structure and spectroscopic analysis of the ferric uptake regulator. *Molecular Microbiology,* 47(4), pp.903-915.

Pollack, J. and Neilands, J. 1970. Enterobactin an iron transport compound from *Salmonella typhimurium*. *Biochemical and Biophysical Research Communications,* 38(5), pp.989-992.

Poole, P. and Allaway, D. 2000. Carbon and nitrogen metabolism in rhizobium. *Advances in Microbial Physiology,* 43, pp.117-163.

Postle, K., Kastead, K.A., Gresock, M.G., Ghosh, J. and Swayne, C.D. 2010. The TonB dimeric crystal structures do not exist *in vivo*. *Mbio,* 1(5), pp.e00307.

Prabhu, V., Biolchini, P. and Boyer, G. 1996. Detection and identification of ferricrocin produced by ectendomycorrhizal fungi in the *Genus wilcoxina*. *Biometals,* 9(3), pp.229-234.

Qi, Z., Hamza, I. and O'Brian, M. 1999. Heme is an effector molecule for iron-dependent degradation of the bacterial iron response regulator (Irr) protein. *Proceedings of the National Academy of Sciences of the United States of America,* 96(23), pp.13056-13061.

Qi, Z. and O'Brian, M. 2002. Interaction between the bacterial iron response regulator and ferrochelatase mediates genetic control of heme biosynthesis. *Molecular Cell,* 9(1), pp.155-162.

Quandt, J. and Hynes, M.F. 1993. Versatile Suicide Vectors which Allow Direct Selection for Gene Replacement in Gram-Negative Bacteria. *Gene,* 127(1), pp.15-21.

Raymond, K., Dertz, E. and Kim, S. 2003. Enterobactin: An archetype for microbial iron transport. *Proceedings of the National Academy of Sciences of the United States of America,* 100(7), pp.3584-3588.

Reeder, T. and Schleif, R. 1993. AraC protein can activate transcription from only one position and when pointed in only one direction. *Journal of Molecular Biology,* 231(2), pp.205-218.

Renshaw, J., Robson, G., Trinci, A., Wiebe, M., Livens, F., Collison, D. and Taylor, R. 2002. Fungal siderophores: Structures, functions and applications. *Mycological Research,* 106 pp.1123-1142.

Robertson, J. and Lyttleton, P. 1984. Division of peribacteroid membranes in root-nodules of white clover. *Journal of Cell Science,* 69(7), pp.147-157.

Rudolph, G., Semini, G., Hauser, F., Lindemann, A., Friberg, M., Hennecke, H. and Fischer, H. 2006. The iron control element, acting in positive and negative control of iron-regulated *Bradyrhizobium japonicum* genes, is a target for the Irr protein. *Journal of Bacteriology,* 188(2), pp.733-744.

Sandy, M. and Butler, A. 2009. Microbial iron acquisition: Marine and terrestrial siderophores. *Chemical Reviews,* 109(10), pp.4580-4595.

Sangwan, I., Small, S.K. and O'Brian, M.R. 2008. The *Bradyrhizobium japonicum* Irr protein is a transcriptional repressor with high-affinity DNA-binding activity. *Journal of Bacteriology,* 190(15), pp.5172-5177.

Schalk, I.J. and Guillon, L. 2013. Fate of ferrisiderophores after import across bacterial outer membranes: Different iron release strategies are observed in the cytoplasm or periplasm depending on the siderophore pathways. *Amino Acids,* 44(5), pp.1267-1277.

Schleif, R. 2010. AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMS Microbiology Reviews,* 34(5), pp.779-796.

Schwyn, B. and Neilands, J. 1987. Universal chemical-assay for the detection and determination of siderophores. *Analytical Biochemistry,* 160(1), pp.47-56.

Shea, C. and McIntosh, M. 1991. Nucleotide-sequence and genetic organization of the ferric enterobactin transport-system - homology to other periplasmic binding protein-dependent systems in *Escherichia coli*. *Molecular Microbiology,* 5(6), pp.1415-1428.

Sidebottom, A.M., Johnson, A.R., Karty, J.A., Trader, D.J. and Carlson, E.E. 2013. Integrated metabolomics approach facilitates discovery of an unpredicted natural product suite from *Streptomyces coelicolor* M145. *Acs Chemical Biology,* 8(9), pp.2009-2016.

Singleton, C., White, G.F., Todd, J.D., Marritt, S.J., Cheesman, M.R., Johnston, A.W.B. and Le Brun, N.E. 2010. Heme-responsive DNA binding by the global iron regulator Irr from *Rhizobium leguminosarum*. *Journal of Biological Chemistry,* 285(21), pp.16023-16031.

Small, S.K., Puri, S., Sangwan, I. and O'Brian, M.R. 2009. Positive control of ferric siderophore receptor gene expression by the Irr protein in Bradyrhizobium japonicum. *Journal of Bacteriology,* 191(5), pp.1361-1368.

Smith, R., Loo, J., Edmonds, C., Barinaga, C. and Udseth, H. 1990. New developments in biochemical mass-spectrometry - electrospray ionization. *Analytical Chemistry,* 62(9), pp.882-899.

Smith, M.J. and Neilands, J.B., (1987). Rhizobactin, a structurally novel siderophore biochemically related to the opines. *In*: *Molecular Strategies for Crop Protection*, pp.157-168. Liss, A.R., Inc.

Stancheva I., Geneva M., Djonova E., Kaloyanova N., Sichanova M., Boychinova M. and Georgiev G. 2008. Response of alfafa (*Medicago sativa*) growth at low accessible phosphorus source to the dual inoculation with mycorrhizal fungi and nitrogen fixing bacteria. *General and Applied Plant Physiology,* 34(3-4), pp.319-326.

Stevens, J., Carter, R., Hussain, H., Carson, K., Dilworth, M. and Johnston, A. 1999. The *fhu* genes of *Rhizobium leguminosarum*, specifying siderophore uptake proteins: FhuDCB are adjacent to a pseudogene version of *fhuA*. *Microbiology-Uk,* 145, pp.593-601.

Storey, E.P., Boghozian, R., Little, J.L., Lowman, D.W. and Chakraborty, R. 2006. Characterization of 'schizokinen'; a dihydroxamate-type siderophore produced by *Rhizobium leguminosarum* IARI 917. *Biometals,* 19(6), pp.637-649.

Todd, J. D., Wexler, M., Sawers, G., Yeoman, K., Poole, P. and Johnston, A. 2002. RirA, an iron-responsive regulator in the symbiotic bacterium *Rhizobium leguminosarum*. *Microbiology-Sgm,* 148pp.4059-4071.

Todd, J.D., Sawers, G., Rodionov, D.A. and Johnston, A.W.B. 2006. The *Rhizobium leguminosarum* regulator IrrA affects the transcription of a wide range of genes in response to Fe availability. *Molecular Genetics and Genomics,* 275(6), pp.564-577.

Traxler, M.F., Watrous, J.D., Alexandrov, T., Dorrestein, P.C. and Kolter, R. 2013. Interspecies interactions stimulate diversification of the *Streptomyces coelicolor* secreted metabolome. *Mbio,* 4(4), pp.e00459-13.

Vanderpool, C. and Armstrong, S. 2004. Integration of environmental signals controls expression of *Bordetella* heme utilization genes. *Journal of Bacteriology,* 186(4), pp.938-948.

Vasil, M. and Ochsner, U. 1999. The response of *Pseudomonas aeruginosa* to iron: Genetics, biochemistry and virulence. *Molecular Microbiology,* 34(3), pp.399-413.

Viguier, C. 2005. Genetic regulation of Iron Responsive Genes in *Sinorhizobium meliloti*. PhD Thesis. Dublin City University.

Viguier, C., O Cuiv, P., Clarke, P. and O'Connell, M. 2005. RirA is the iron response regulator of the rhizobactin 1021 biosynthesis and transport genes in *Sinorhizobium meliloti 2011*. *FEMS Microbiology Letters,* 246(2), pp.235-242.

Waring, W.S. and Werkman, C.H., (1942). Growth of bacteria in an iron-free medium. *Arch. Biochem.* 1, pp.303-310.

Wexler, M., Todd, J.D., Kolade, O., Bellini, D., Hemmings, A.M., Sawers, G. and Johnston, A.W. 2003. Fur is not the global regulator of iron uptake genes in *Rhizobium leguminosarum*. *Microbiology-Sgm,* 149(5), pp.1357-1365.

Wexler, M., Yeoman, K., Stevens, J., de Luca, N., Sawers, G. and Johnston, A. 2001. The *Rhizobium leguminosarum tonB* gene is required for the uptake of siderophore and haem as sources of iron. *Molecular Microbiology,* 41(4), pp.801-816.

White, G.F., Singleton, C., Todd, J.D., Cheesman, M.R., Johnston, A.W.B. and Le Brun, N.E. 2011. Heme binding to the second, lower-affinity site of the global iron regulator Irr from *Rhizobium leguminosarum* promotes oligomerization. *Febs Journal,* 278(12), pp.2011-2021.

Wickstrum, J.R., Skredenske, J.M., Kolin, A., Jin, D.J., Fang, J. and Egan, S.M. 2007. Transcription activation by the DNA-binding domain of the AraC family protein RhaS in the absence of its effector-binding domain. *Journal of Bacteriology,* 189(14), pp.4984-4993.

Wiener, M., Freymann, D., Ghosh, P. and Stroud, R. 1997. Crystal structure of colicin ia. *Nature,* 385(6615), pp.461-464.

Winkelmann, G. 1990. Structural and stereochemical aspects of iron transport in fungi. *Biotechnology Advances,* 8(1), pp.207-231.

Winkelmann, G. 2007. Ecology of siderophores with special reference to the fungi. *Biometals,* 20(3-4), pp.379-392.

Winterberg, B., Uhlmann, S., Linne, U., Lessing, F., Marahiel, M.A., Eichhorn, H., Kahmann, R. and Schirawski, J. 2010. Elucidation of the complete ferrichrome A biosynthetic pathway in *Ustilago maydis*. *Molecular Microbiology,* 75(5), pp.1260-1271.

Wyckoff, E.E., Mey, A.R. and Payne, S.M. 2007. Iron acquisition in *Vibrio cholerae*. *Biometals,* 20(3-4), pp.405-416.

Wyckoff, E.E. and Payne, S.M. 2011. The *Vibrio cholerae* VctPDGC system transports catechol siderophores and a siderophore-free iron ligand. *Molecular Microbiology,* 81(6), pp.1446-1458.

Yamamato, S., Ikai, H., Uesugi, T., Horie, A. and Hirai, Y. 1995. Occurrence and antigenic heterogeneity of L-2,4-diaminobutyrate decarboxylase in *Acinetobacter* species. *Biological & Pharmaceutical Bulletin,* 18(3), pp.454-456.

Yamanaka, K., Oikawa, H., Ogawa, H., Hosono, K., Shinmachi, F., Takano, H., Sakuda, S., Beppu, T. and Ueda, K. 2005. Desferrioxamine E produced by *Streptomyces griseus* stimulates growth and development of *Streptomyces tanashiensis*. *Microbiology-Sgm,* 151 pp.2899-2905.

Yang, J., Ishimori, K. and O'Brian, M. 2005. Two heme binding sites are involved in the regulated degradation of the bacterial iron response regulator (Irr) protein. *Journal of Biological Chemistry,* 280(9), pp.7671-7676.

Yang, J., Sangwan, I., Lindemann, A., Hauser, F., Hennecke, H., Fischer, H. and O'Brian, M. 2006. *Bradyrhizobium japonicum* senses iron through the status of haem to regulate iron homeostasis and metabolism. *Molecular Microbiology,* 60(2), pp.427-437.

Yang, J., Sangwan, I. and O'Brian, M.R. 2006. The *Bradyrhizobium japonicum* Fur protein is an iron-responsive regulator *in vivo*. *Molecular Genetics and Genomics,* 276(6), pp.555-564.

Yeoman, K., Curson, A., Todd, J., Sawers, G. and Johnston, A. 2004. Evidence that the rhizobium regulatory protein RirA binds to cis-acting iron-responsive operators (IROs) at promoters of some Fe-regulated genes. *Microbiology-Sgm,* 150, pp.4065-4074.

Yeoman, K., Wisniewski-Dye, F., Timony, C., Stevens, J., deLuca, N., Downie, J. and Johnston, A. 2000. Analysis of the *rhizobium leguminosarum* siderophore-uptake gene *fhuA*: Differential expression in free-living bacteria and nitrogen-fixing bacteroids and distribution of an *fhuA* pseudogene in different strains. *Microbiology-Uk,* 146, pp.829-837.

Yuan, W., Gentil, G., Budde, A. and Leong, S. 2001. Characterization of the *Ustilago maydis sid2* gene, encoding a multidomain peptide synthetase in the ferrichrome biosynthetic gene cluster. *Journal of Bacteriology,* 183(13), pp.4040-4051.

Yue, W., Grizot, S. and Buchanan, S. 2003. Structural evidence for iron-free citrate and ferric citrate binding to the TonB-dependent outer membrane transporter FecA. *Journal of Molecular Biology,* 332(2), pp.353-368.