From prediction to function: Polyamine biosynthesis and formate metabolism in the α- and ε-Proteobacteria

A thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

Frances Louise Shaw

Institute of Food Research
Norwich Research Park
Norwich NR4 7UA

June 2011

To Chris

for the past
for the future
for everything

Acknowledgements

Funding for this project was provided by BBSRC through a doctoral training grant awarded to the IFR and it is my pleasure to thank those who made this thesis possible. I thank my wonderful friends and family, both near and far, my supervisors, advisors, colleagues, group members and coffee-time friends, all whom have kept me as sane as is necessary over the past four years.

In particular, I am thankful to my supervisor, Dr Arnoud van Vliet, whose encouragement, supervision, patience and support has been always present; from my adoption into the Campy Group, to the submission day. I am grateful for Arnoud's blackberry which has made it possible for him to contact me at all times.

Thank-you to everyone who has contributed to the data and helped with the experimentation in this thesis: Matthew Burrell, Julea Butt, Allan Downie, Anne Edwards, Kath Elliott, Christine Fuell, Duncan Gaskin, Colin Hanfrey, Andy Hitchcock, Dave Kelly, Gwen Le Gall, Sophie Marritt, Tony Michael, Fran Mulholland, Bruce Pearson, Mark Reuter & Arnoud van Vliet.

I owe my deepest gratitude to my group members, both past and present, who have made the days in the lab entertaining, enjoyable and fun. Arnoud, Bruce, Christine, Colin, Don, Duncan, Faye, Fran, Helen, Kath, Louise, Lynn, Maddy, Mark, Matt, Neil, Rebecca, Rob, Thanh & Tony, I extend my appreciation to you all. My friendship with Christina & Rosy has been a blessing and a delight: thank goodness we had so many weddings to plan as distractions from science.

I thank God for giving me the intellect to understand the complexity of the project, the strength to complete this research, and for sending his angels when I needed them the most. Lastly, I offer my kindest regards to everyone who supported me, in any respect, during the completion of my PhD.

Ecclesiastes Chp 1, v 18.

Abstract

Understanding gene function is a fundamental requirement in the comprehension of diversity and evolution of bacterial species, and elucidation of how microbial genetic material is ordered and regulated. This investigation has highlighted the function of genes previously un-annotated or uncharacterised. Studies such as this are essential because the ability to sequence genomes is quicker, cheaper and faster than ever before and genomic data are being generated at an exponential rate. Genome annotation relies heavily on the comparison of predicted open reading frames with previously characterised genes. Therefore, the need to comprehend these data is more important than ever before. If we are to gain a complete understanding of biological systems, we need to exploit the synergistic relationship between basic experimental biology and computer generated genome annotation. At present, while genome annotation can suggest gene function, it lacks the gravitas that only experimental approaches can bring. The conclusions of this thesis demonstrate, using two case studies, how both experimental biology and existing genome annotation can be used to enhance our comprehension of genome biology. Firstly, the diversity of polyamine biosynthesis, focusing on the occurrence and significance of a triamine, homospermidine, has been explored. It shows that the specific nature, size and structure of the triamine is not crucial for normal cell growth in various model micro-organisms, suggesting that the divergence in polyamine profile in these species has perhaps occurred by chance, rather than directed evolution. In addition, it suggests that the homospermidine synthase gene has evolved vertically primarily in the α-Proteobacteria and by horizontal gene transfer to other bacteria, bacteriophage, archaea, eukaryotes and viruses. Secondly, using a suite of laboratory methods (proteomics, transcriptomics and metabolomics) and bioinformatic investigation, defined roles have been assigned to previously uncharacterised or putatively annotated genes involved with the production of a selenoprotein, formate dehydrogenase, in the human pathogen *Campylobacter jejuni*.

Declaration

I hereby declare that this thesis and the work described herein are original, except where indicated by reference or otherwise, and has not previously been submitted for any degree at this or any other university.

Signed: ……..

F L Shaw

Date: ……..

Parts of the research described in this thesis have appeared in the scientific literature as follows:

Frances L. Shaw, Katherine. A. Elliott, Lisa N. Kinch, Christine Fuell, Margaret A. Phillips and Anthony J. Michael (2010). "Evolution and multifarious horizontal transfer of an alternative biosynthetic pathway for the alternative polyamine sym-homospermidine." Journal of Biological Chemistry **285**(19): 14711-23.

Table of Contents

TABLE OF FIGURES

TABLE OF TABLES

Collaborative Work

Chapter Four

Phylogenetic tree and alignment construction performed by Tony Michael (Figure 4.5 and Appendices I & II)

The cloning and expression of the genes and the HPLC analysis was performed by Kath Elliott and myself (Figure 4.6)

LC-MS/MS performed by Christine Fuell (Figure 4.7)

Thanks to Matthew Burrell for his help with the enzymatic characterisation work.

Chapter Five

C. jejuni mutagensis performed by Bruce Pearson, and phenotype analysis was performed by Colin Hanfrey (Figures 5.11-13)

Thanks to Allan Downie & Anne Edwards for their help with the Tn5 transposon insertion library.

Chapter Eight

NMR analysis by Gwen Le Gall (Figure 8.14)

Thanks to Bruce Pearson with his help in interpreting the microarray data.

Thanks to Fran Mulholland for the spot pick and protein identification and for help with the proteomics methods.

Chapter Nine

Protein-protein interaction and network analysis performed with help from Mark Reuter (Figures 9.14)

Chapter One

Thesis Introduction

TABLE OF FIGURES

Chapter One - Thesis Introduction

Background Information

Technical advances such as the development of molecular cloning, next generation sequencing, PCR and oligonucleotide microarrays are essential to our current capacity to sequence, annotate and study complete genomes. Recent years have seen the development of a variety of so-called 'next-generation' sequencing platforms, with several others anticipated to become available shortly [1]. Genome sequencing is easier, cheaper and faster than ever before, but the annotation of these genomes is a tedious task, normally performed by automated methods. Poptsova [2] has described some of the problems arising from mis-annotation of microbial genomes. The accuracy of these automated-annotation methods has been questioned since the beginning of the sequencing era. Genome annotation is a multilevel process, and errors can emerge at different stages: during sequencing, as a result of gene-calling procedures and in the process of assigning gene functions. Missed or wrongly annotated genes differentially impact various types of analyses. Researchers need to be aware of the existing errors in the annotation of even well-studied genomes, such as *Escherichia coli*, and consider additional quality controls for their results [2]. Genes are annotated based on sequence similarity to other previously annotated genes. However, it is plausible that if a gene was annotated based on sequence similarity to a second gene, and if this second gene was annotated based on sequence similarity to a third, then it is possible that ultimately one gene can be annotated as something, which is has little or no similarity to. This relationship of possible mis-annotation is shown in Figure 1.1.

There are several methods and tools available to assemble, map and interpret huge quantities of relatively or extremely short nucleotide sequence data. These approaches, which have been introduced for functional genomic investigation of next-generation sequencing data, have been reviewed [1, 3]. However, ignoring, or not updating gene annotations, the fundamental basis for any bioinformatic investigation can lead to a magnitude of problems for researchers trying to understand bacterial genomes. The only method of unequivocally determining the function of a gene is by either recombinant

expression of the gene in an appropriate host strain, or genetic manipulation of the gene, followed by phenotypic testing.

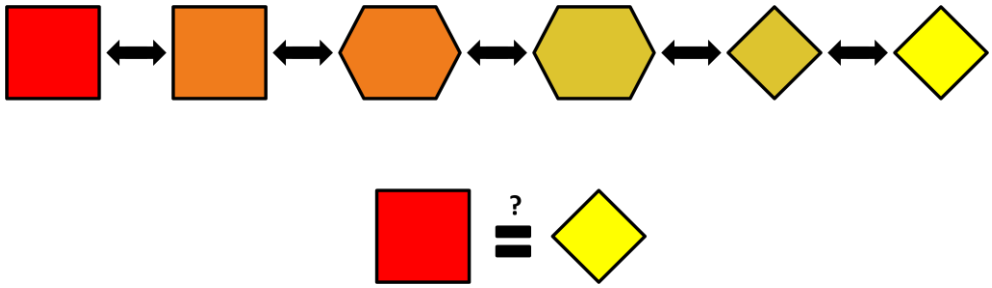


Figure 1.1 The possible mis-annotation of genes based on multiple annotations with similarity to one another.
The gene on the left (red square) has been biochemically characterised and named accordingly. The gene on the right (yellow diamond) may have been annotated the same as red square after a progression of annotations, when in actual fact the genes have little or no similarity to one another.

There are several approaches to handling these large genomic sequence data sets to enable accurate annotation and interpretation. A summary of some of these approaches are detailed here:

Proteogenomics is a method which combines genomics and proteomics; high-throughput identification and characterisation of proteins by extra-large shotgun MS/MS approaches and the integration of these data with genomic data [4]. Armengaud [5] describes how intimately proteogenomics and systems biology are related to one another, and high-quality genome annotation is a crucial basis for this method. This method can be used in the discovery of remaining un-annotated genes, defining translational start sites, listing signal peptide processing events and post-translational modifications [4, 5]. The proteomics and genomics alliance produces almost complete and accurate gene catalogues for small microbial genomes, a comprehensiveness which is essential for efficient systems biology [6]. However, expression of a particular gene does not always result in a corresponding protein function. Specific understanding of particular gene regulation pathways is necessary for complete comprehension.

The Gene Ontology (GO) is a standardised system for the capture of genomic data to ensure that the information produced from the ever-increasing

number of microbial sequencing projects can be effectively utilised. The GO provides the standard for gene product annotations in the areas of molecular function, biological process and cellular component. Since it began, there have been changes and additions to the GO annotation format and evidence storage system to reflect the needs of the microbial annotation community. The capture of annotation information with systems like the GO is absolutely essential to enable the efficient mining of annotation information across diverse genomes and thus to further biological research in meaningful ways [7].

Hrmova and Fincher [8] have reviewed the subject of how structural biology can be integrated with and adds value to functional genomics programs. Structural biology can assist in the definition of gene function through the identification of the likely function of the protein products of genes. Using the 3D information from DNA sequences, protein sequences can be classified into broad groups, according to the overall 'fold', or 3D shape, of the protein. Structural information can be used to predict the preferred substrate of a protein, and thereby greatly enhance the accurate annotation of the corresponding gene. Furthermore, it can enable the effects of amino acid substitutions in enzymes to be better understood with respect to enzyme function and could thereby provide insights into natural variation in genes [8]. However, as with all bioinformatic-only analyses, the results are limited by the quality and accuracy of the sequencing data. The exact function of a gene, can be 'predicted' using its DNA sequence and corresponding 3D information, however the function can only be unequivocally tested with biological manipulation of the organism or gene.

Genome scale metabolic modelling is a process which involves the construction and interpretation of data from genome sequence annotation with biochemical and physiological experimentation. Systematic analysis of biological metabolic processes by means of modelling and simulations has made the identification of metabolic networks and prediction of metabolic capabilities under different conditions possible [9]. Metabolic modelling allows insights into a bacteria, such as the minimum requirement for amino acids and other carbon substrates necessary for sustenance [10]. The data generated from gene deletion studies is essential for the completion of

genome-wide metabolic modelling as it is important to determine the absolute requirement of genes, or if alternative biosynthetic and metabolic pathways are available. Secondly, robust annotation of genes is crucial for these studies to be accurate interpretations of the metabolism of an organism. The results generated in these studies have shown that the combination of *in silico* modelling with experimental technologies is an effective strategy to enhance biological discovery for less characterised organisms and their genomes. The rapid progress of molecular biology tools for directed genetic modifications, accurate quantitative experimental approaches, high-throughput measurements, together with development of genome sequencing has made the foundation for a new area of metabolic engineering that is driven by metabolic models [9].

Thesis Aims

- To elucidate gene function of either unannotated or uncharacterised bacterial genes by biological manipulation of appropriate bacterial genomes and organisms.

- To use the information derived from this bacterial manipulation (mentioned above) to determine facts about the metabolism and cellular functions of the organism, or predict how the specific genes have been conserved throughout different species.

The focus of this project has been upon the evolution, phylogenetic distribution and composition of biosynthetic pathways, using different species of bacteria as model organisms. Biochemistry, molecular biology, analytical chemistry, microbiology, proteomics, transcriptomics, metabolomics and comparative and functional genomics have all been used as tools to comprehend the bacterial genome and elucidate specific gene function. This combination approach which uses an amalgamation of comparative genomics and molecular microbiological techniques alongside analytical chemistry allows us to paint a picture of genome evolution and the development of the bacterial biosynthetic pathways.

During this investigation, the universal model micro-organism, *Escherichia coli*, pathogen, *Campylobacter jejuni,* and plant symbiont, *Rhizobium leguminosarum,* have been used as model bacterial organisms. Ultimately, the evolution and diversity of bacterial pathways has been investigated, with particular attention and focus on a small protein with a crucial role in formate dehydrogenase activity, and the phylogenetically diverse homospermidine synthase enzyme.

The thesis is split into two distinct case studies, both which demonstrate the need for robust genome annotation to further microbial genome understanding.

Section One – Polyamines and Gene Conservation

The project explored the diversity of polyamine biosynthesis, focusing on the occurrence and significance of the triamine, homospermidine. A homo-spermidine-synthetic enzyme (HSS) from *Bradyrhizobium japonicum* has been recombinantly expressed and purified and kinetically characterised. The kinetic parameters of the *B. japonicum* HSS were similar to those from previously characterised HSS enzymes. It has been demonstrated that the HSS gene has evolved vertically primarily in the α-Proteobacteria and by horizontal gene transfer to other bacteria, bacteriophage, archaea, eukaryotes and viruses. This phenonomen was demonstrated when enzymatically active, diverse HSS orthologues were identified by BLAST searches and biochemically tested for their ability to produce homospermidine. It has been shown that *E. coli* can produce homospermidine and other non-native polyamine products when expressing HSS orthologues. In addition, it has been shown that homospermidine (or an alternative non-native triamine) is required for normal growth of the α-Proteobacterium *R. leguminosarum,* incidentally demonstrating for the first time that the clinically approved cancer drug α-difluoromethylornithine (DFMO) has anti-bacterial activity. The specific nature, size and structure of the triamine was investigated in *R. leguminosarum* and *C. jejuni* and it was found that the exact length of the triamine was not crucial for normal cell growth. The findings in this investigation have provided insight into the crucial, but perhaps unspecific, role of polyamines in micro-organisms.

Section Two – *Campylobacter jejuni* and Selenoproteins

The role of one particular protein, previously identified by consistent up-regulation during acid shock of *C. jejuni*, has been studied by inactivation of the gene by an insertion antibiotic cassette. Interestingly, the insertion activation had no effect on the growth, motility, auto-agglutination, resistance to acid shock or aerotaxis of the cells, but did however remove nearly all formate dehydrogenase (FDH) activity. Proteomics, whole genome microarray and metabolomics were used to further define the role of the protein, and a bioinformatic investigation indicated that the protein is most likely involved with selenium up-take. FDH is a selenoprotein and activity could be restored in the insertion inactivation mutant cells by the re-expression of the mutated gene, or the addition of exogenous selenium. The pathway of selenium up-take and FDH synthesis in *C. jejuni* have been described in this investigation based on previous gene annotation, biochemical experimentation and the available literature. FDH and selenoprotein synthesis has never been described in *C. jejuni* previously. The findings in this investigation have interesting implications both in the field of *C. jejuni* metabolism and bacterial selenoprotein synthesis.

Impact and importance of this investigation

Investigations such as this, which piece together many aspects of biology are increasingly necessary. As the number of sequenced genomes increases, it is important that the annotation of these genomes are continually examined and tested. Mis-annotation is a common occurrence and the true mechanism or function of a gene can only be determined by experimental testing.

Additional Information

In between July – December 2009, the Institute of Food Research, Norwich, underwent a structural reorganisation and the Microbiome & Metabolism Group where the PhD project began (October 2007) was no longer in line with the strategic vision of the Institute. The projects primary supervisor, Dr.

Anthony Michael, left the Institute, as did all the other group members. 'Section One' of the thesis was executed under Dr. Michael's supervision. In compliance with the Institute's plan, this research could not be continued.

During the third year of the PhD, the project moved into the Foodborne Bacterial Pathogens programme (January 2010) within the Institute. The remainder of the PhD project was completed in Dr. Arnoud van Vliet's group and the research is described in 'Section Two'. Consequently, you will notice that the thesis has been divided into two distinct sections.

Most of the research under Dr. Michael's supervision (Chapters Four & Five) was published in 'Evolution and multifarious horizontal transfer of an alternative biosynthetic pathway for the alternative polyamine sym-homospermidine', *Journal of Biological Chemistry.* March 1, 2010. 285(19): 14711-23. This paper was selected by the Associate Editors and Editorial Board Members as a 'Paper of the Week.' The Papers of the Week highlight the research papers the JBC consider to be the best they receive for publication. They represent the top 1% of papers reviewed in terms of significance and overall importance. This work was also presented orally and in poster form at the Gordon Research Conference – Polyamines, New Hampshire, USA. June 2008.

The research in the Foodborne Bacterial Pathogens Programme was presented orally at the Society for General Microbiology Conference, Nottingham, UK; September 2010, and the aim is to submit a manuscript about this work in 2011.

Chapter One – References

1. Horner DS, Pavesi G, Castrignano T, De Meo PD, Liuni S, Sammeth M, Picardi E, Pesole G: **Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing**. *Brief Bioinformatics* 2010, **11**(2):181-197.
2. Poptsova MS, Gogarten JP: **Using comparative genome analysis to identify problems in annotated microbial genomes**. *Microbiology* 2010, **156**(7):1909-1917.
3. Ussery D, Wassenaar T, Borini S: **Computing for Comparative Microbial Genomics - Bioinformatics for Microbiologists**, 1 edn. London: Springer; 2009.
4. Castellana N, Bafna V: **Proteogenomics to discover the full coding content of genomes: a computational perspective**. *Journal of Proteomics* 2010, **73**(11):2124-2135.
5. Armengaud J: **Proteogenomics and systems biology: quest for the ultimate missing parts**. *Expert Reviews in Proteomics* 2010, **7**(1):65-77.
6. Armengaud J: **A perfect genome annotation is within reach with the proteomics and genomics alliance**. *Current Opinions in Microbiology* 2009, **12**(3):292-300.
7. Giglio MG, Collmer CW, Lomax J, Ireland A: **Applying the Gene Ontology in microbial annotation**. *Trends in Microbiology* 2009, **17**(7):262-268.
8. Hrmova M, Fincher GB: **Functional genomics and structural biology in the definition of gene function**. *Methods in Molecular Biology* 2009, **513**:199-227.
9. Cvijovic M, Olivares-Hernandez R, Agren R, Dahr N, Vongsangnak W, Nookaew I, Patil KR, Nielsen J: **BioMet Toolbox: genome-wide analysis of metabolism**. *Nucleic Acids Research* 2010, **38**:144-149.
10. Schilling CH, Covert MW, Famili I, Church GM, Edwards JS, Palsson BO: **Genome-scale metabolic model of *Helicobacter pylori* 26695**. *Journal of Bacteriology* 2002, **184**(16):4582-4593.

Chapter Two

Materials & Methods

TABLE OF TABLES

Chapter Two - Materials & Methods

2.1 General Lab Methods

2.1.1 Chemicals

All chemicals were purchased from Sigma-Aldrich, unless stated otherwise. Ultrapure water was used throughout and was purified using a MilliQ water system (Barnstead).

Homospermidine is not available commercially, and was obtained from Jeongmi Lee of the Dept. of Pharmacology, UT Southwestern Medical Center, Dallas, USA.

Restriction enzymes were purchased either from New England Biolabs or Promega.

2.1.2 Autoclaving and Sterilisation

Solutions, pipette tips and bacterial growth media were sterilised by autoclaving at 69 KPa for 20 minutes in a B&T Autoclave 225 5H (Laboratory Thermal Equipment). When autoclaving was not appropriate, solutions were filter-sterilised with 0.2 μM Ministart filter unit (Sartorius Stedim Biotech).

2.1.3 Aseptic Technique

When working with bacterial cultures, standard aseptic techniques were adhered to using a bunseon burner flame or a class II microbiological safety cabinet BS 5726, Part1:1992 (Walker).

2.2 Bacterial Strains Used

Table 2.1 summarises the bacterial strains used in this investigation.

Strain	Genotype	Source
Escherichia coli Top10	-	Invitrogen
Escherichia coli XL2 Blue	-	Stratagene
Escherichia coli BL21(DE3)	-	Novagen
Campylobacter jejuni 81116 (NCTC11828)	Wild-type	routinely used at the IFR
Campylobacter jejuni NCTC 11168	Wild-type	routinely used at the IFR
Bradyrhizobium japonicum N2P5549	Wild-type	kindly supplied by Professor Allan Downie (Department of Molecular Microbiology, John Innes Centre, Norwich, UK).
Rhizobium leguminosarum bv. viciae 3841	Wild-type	kindly supplied by Professor Allan Downie (Department of Molecular Microbiology, John Innes Centre, Norwich, UK).

Table 2.1 Bacterial strains used in this thesis.

2.3 Kits Used

Agilent RNA 6000 Nano kit
Agilent Hi-RPM Gene Expression Hybridization Kit
Amersham 2D Protein Quant Kit
Sigma midi-prep
Merck OnEx Overnight Protein Auto-induction
Novagen Bugbuster
Novagen T7-tag protein purification system Novagen
Promega pGEM-T-EASY vector system
Qiagen QiaexII Gel Purification
Qiagen PCR Clean-up Reaction
Qiagen DNAeasy
Qiagen RNAeasy
Qiagen mini-prep
Qiagen mini-spin-prep
Sigma Ammonia Assay Kit

2.4 Media

All media were prepared by adding ingredients to 1 litre of water and autoclaving at 69 KPa for 20 minutes. For plates 15 g Difco agar, or 8 g agarose (polyamine free) was added. Antibiotics, XGal or IPTG were added aseptically after auto-claving when necessary.

2.4.1 LB Medium

10 g Bacto™ tryptone (Difco Laboratories)
5 g Bacto™ yeast extract (Difco Laboratories)
10 g NaCl
pH to 7.0 with NaOH

2.4.2 TY Medium

5 g Bacto™ tryptone (Difco Laboratories)
3 g Bacto™ yeast extract (Difco Laboratories)
1.4 g $CaCl_2 \cdot 6H_2O$

2.4.3 Polyamine-free Minimal Medium

Sigma D5030-1 L powder

3.7 g Sodium bicarbonate

0.5 g Aspartic acid

0.5 g Serine

0.2 g Sodium pyruvate

0.1 g Cysteine

0.5 g Glutamic acid

0.1 g Proline

5 mg $FeCl_3 \cdot 6H_2O$

10 mg Adenine sulphate

2.4.4 Acid Minimal Salts Medium (AMS)

0.5 ml 1 M K_2HPO_4

0.5 g $MgSO_4 \cdot 7H_2O$

0.2 g NaCl

4.19 g Mops

1 ml Solution A (15 g EDTA-Na_2, 0.16 g $ZnSO_4 \cdot 7H_2O$, 0.2 g $NaMoO_4 \cdot 2H_2O$, 0.25 g Boric Acid, 0.2 g $MnSO_4 \cdot 4H_2O$, 0.02 g $CuSO_4 \cdot 5H_2O$, 1 mg $CoCl_2 \cdot 6H_2O$, made up to 1 l in de-ionised water)

2 ml Solution B (1.28 g $CaCl_2 \cdot 2H_2O$, 0.33 g $FeSO_4 \cdot 7H_2O$ made up to 100 ml in de-ionised water)

1 ml Solution C (1 g Thiamine HCl, 2 g D-pantothenic acid calcium salt, 1 mg Biotin, made up to 1 l in de-ionised water)

Solutions A, B & C were filter sterlised then added aseptically after the medium was auto-claved.

2.4.5 Brucella Medium

10 g Bacto™ Pancreatic digest of casein (Difco Laboratories)

10 g Bacto™ Peptic digest of Animal Tissue (Difco Laboratories)

1 g Dextrose

5 g Yeast extract

5 g NaCl

0.1 g Sodium bisulfite

2.4.6 Skirrow Medium (recipe from Blood agar base No.2, Oxoid Ltd.)

15 g Proteose peptone

2.5 g Liver Digest

5.0 g Yeast Extract

5.0 g NaCl

pH 7.4

Campylobacter selective supplement (Skirrow, SR0069E, Oxoid Ltd. Final conc. Vancomycin 0.1 $mg.ml^{-1}$, Trimethoprim 50 $\mu g.ml^{-1}$, Polmyxin B 2500$IU.L^{-1}$) was added after auto-claving.

2.4.7 Antibiotics, X-gal and IPTG used in Media

Antibiotics were made to concentrated solution in water or water:ethanol (50:50) and filter sterilised (0.2 μM Sartorius Stedim Biotech) then added aseptically to the medium after auto-claving.

X-gal (bromo-chloro-indolyl-galactopyranoside) was made into solution with dimethyl formamide and added aseptically to the medium after auto-claving.

IPTG (Isopropyl β-D-1-thiogalactopyranoside) was dissolved with water and added aseptically to the medium after auto-claving.

2.5 Bacterial Growth

Rhizobium leguminosarum and *Bradyrhizobium japonicum* were grown in AMS or TY liquid broth at 28°C with constant agitation at 200 rpm. Growth was monitored by optical density at 600 nm. AMS was used for cultures in polyamine experiments as TY contains trace amounts of polyamine. The cells were harvested after four to six days. *R. leguminosarum* and *B. japonicum* were grown on solid TY plates made with agar, or solid AMS plates made with agarose, at 28°C unless stated otherwise.

Escherichia coli (all strains) were grown in LB liquid broth at 37°C with constant agitation at 200 rpm. Growth was monitored by optical density at 600 nm. Cells were harvested after several hours of growth, depending on the experiment. *E. coli* (all strains) were grown on LB plates, with or without appropriate additives depending on the experiment, at 37°C unless stated otherwise.

Campylobacter jejuni (all strains) were grown in Brucella broth at 37°C in the microaerobic cabinet (MACS-MG-1000 Don Whitley Scientific) under microaerobic conditions (85% N_2, 5% O_2, 10% CO_2) with constant agitation at 200 rpm. Cells were harvested after several hours of growth, depending on the experiment. *C. jejuni* (all strains) were grown on Brucella or Skirrow plates, with or without appropriate additives depending on the experiment, at 37°C also in the microaerobic cabinet. When doing polyamine experiments in *C. jejuni,* Polyamine-free Minimal Media was used.

2.5.1 Bacterial Glycerol Stocks

E. coli (all strains) were grown overnight in LB liquid cultures (10 ml) and harvested by centrifugation (20,000 x *g*, 10 minutes) and resuspended in LB with 20% glycerol for long term storage at -80°C. *C. jejuni* (all strains) were grown overnight in Brucella liquid cultures (50 ml) and harvested by centrifugation (10,000 x *g*, 10 minutes) and then resuspended in bead stocks (Technical Service Consultants Ltd) for long term storage at -80°C.

Glycerol stocks were made of all mutants, plasmids, strains and construct combinations.

2.6 Molecular Biology

2.6.1 Agarose Gels

Gels were prepared for the visualisation of DNA fragments using 0.8-2.0% (w/v) (depending on fragment size) agarose in 1 x TAE or TBE buffer containing 0.0002 vol SafeView staining dye (NBS Biological). DNA was loaded with 6 x bromophenol blue loading dye (0.25% bromophenol blue,

50% glycerol) and run for approximately 30 minutes at 70 V (mini gel – Horison 58, Gibco BRL) or 45 minutes at 100 V (larger gels – Horizon 11.14, Gibco BRL). The gels were visualised using ultra-violet radiation (300 nm) and photographed using U:Genius (Syngene) gel doc system.

2.6.2 Preparation of cellular genomic DNA

***Bradyrhizobium japonicum*:** A method developed for genomic DNA preparation from Agrobacteria and Mezorhizobia was applied. The bacteria were grown until stationary phase in TY liquid medium in 50 ml cultures and harvested by centrifugation (4,500 x *g*, 4°C, 15 minutes), washed with TNE buffer (10 mM Tris pH 8, 10 mM EDTA, 10 mM NaCl), then resuspended in TNE + 1% Triton. Lysozyme (10 $mg.ml^{-1}$) was added and incubated at 37°C for 30 minutes, followed by the addition of proteinase K (10 $mg.ml^{-1}$) and a further incubation at 37°C for 1 hour. Two phenol:chloroform:isoamylalcohol (IAA) (25:24:1), followed by a chloroform:IAA (24:1) extractions were performed and the aqueous phase removed. DNA was precipitated with 0.05 vol 5 M NaCl and 3 vol ethanol, and left at -20°C for 1 hours before being harvested by centrifugation (4,500 x *g*, 4°C, 20 minutes), washed with 70% ethanol, air dried, resuspended in sterile water and treated with 5 µl RNase ([DNase free] 10 $\mu g.\mu l^{-1}$) for 1 hour at 37°C.

***Campylobacter jejuni*:** The cells were grown on Skirrow plate for 24 hours were harvested in 1 ml Brucella broth. The cells were centrifuged (4,500 x *g*, 4°C, 5 minutes) and the DNA was prepared using a Qiagen DNeasy kit. Firstly, the cells were resuspended in 180 µl ATL buffer. The sample was vortexed, 15 seconds, and 200 µl AL buffer then 200 µl 100% ethanol were added sequentially with vortexing between stages. The sample was centrifuged through the DNeasy Mini Spin column (6,000 x *g*, 1 minute), and the flow through was discarded. The column was washed with 500 µl AW1 buffer (6,000 x *g*, 1 minute), then 500 µl AW2 buffer (20,000 x *g*, 3 minutes). Elution occurred in 200 µl AE buffer applied twice to the filter membrane (1 minute incubation at room temperature, followed by 6,000 x *g*, 1 minute).

2.6.3 Preparation of RNA from *Campylobacter jejuni*

Cells were grown in 50 ml Brucella broth and monitored by OD_{600nm} (the cells were harvested when the culture reached approximately 0.4). 0.1 vol 5% phenol in ethanol was added to each culture and incubated on ice for 30 minutes, then harvested (4,000 rpm, 15 minutes, 4°C). The cells were then resuspended in 1 ml Tri Reagent (Sigma), 0.2 vol chloroform was added and the samples, vortexed (15 seconds), incubated at room temperature (15 minutes) then centrifuged (12,000 x *g*, 4°C, 15 minutes). The aqueous layer was removed and the RNA was prepared using the QIAgen RNeasy kit. 350 µl RTL buffer (containing 0.01 vol β-mercaptoethanol) was added then vortexed, 250 µl 100% ethanol was added and mixed by 2-3 times gentle inversion. The sample was loaded onto the RNeasy Mini Spin column and centrifuged (8,000 x *g*, 30 seconds), followed by a wash with 700 µl RWI buffer (8,000 x *g*, 15 seconds), then 500 µl RPE buffer (8000 x *g*, 15 seconds), then 500 µl RPE buffer (8,000 x *g*, 2 minutes). After a further 1 minute centrifuge (8,000 x *g*) the sample was eluted in 2 x 35 µl RNase-free water at 60°C (incubation on column for 1 minute at room temperature, 20 000 x *g*, 1 minute).

2.6.4 Quantification of DNA

Isolated DNA was run on an agarose gel along-side a DNA ladder of known concentration (Hyperladder IV, Bioline) for quantification of the sample DNA.

2.6.5 Quantification and Purity Determination of RNA

Firstly, the RNA was subjected to the Turbo DNA-free treatment (Ambion) according to the manufacturer's instructions. After the addition of DNase and the appropriate buffer, the sample was incubated at 37°C for 30 minutes. Then centrifuged (10 000 x *g*, 1.5 minutes) before using the RNA 6000 Nano Kit (Agilent, UK) according to manufacturer's instructions to determine the purity of the RNA.

To determine the concentration of the RNA, the Nanodrop Spectrophotometer NS-1000 (Thermo Scientific) was used to calculate the absorbance of the sample at both 260nm and 280nm. Using a calculation specific to RNA, the concentration of the sample was determined using 1.5 µl sample.

2.6.6 Polymerase Chain Reaction (PCR) Amplification

Takara Ex-Taq (Biowhittaker): Amplification of the target gene or DNA fragment was performed in a Hybaid PCR Sprint Thermocycler (Thermo Scientific) using a single initial denaturing stage of 96°C for 2 minutes, then 35 cycles of 96°C for 30 seconds, 50 - 58°C for 45 seconds and 72°C for 1 minute 30 seconds. Finally a single elongation step was performed at 72°C for 5 minutes. The elongation step was altered upon gene length, typically 1 minute per 1 Kb. The annealing temperature was dependent on the oligonucleotides used.

Phusion Taq (New England Biolabs): Amplification of a target gene or DNA fragment with Phusion Taq required a shorter elongation step as the enzyme is able to synthesise 1 Kb in approximately 30 sec.

Hot Star Taq (Qiagen): This master-mix all-in-one reaction system was used for large amounts of replicates of small volume PCR reactions according to the manufacturer's instructions.

Red Taq (Sigma): Used for screening large numbers of colonies by PCR. The advantage of this taq was that the reaction could be placed into the gel without requiring additional loading dye.

Colony PCR: The template DNA was obtained by boiling a single *E. coli* colony in 100 µl water for 10 minutes, then using 1 µl in the reaction mix (either Red or Hot Star Taq, depending on quantity of reactions).

2.6.7 Digestion of DNA

DNA was digested using the appropriate restriction enzymes and buffers (New England Biolabs) according to the manufacturer's instructions and incubated at 37°C for 30 minutes. Single digestions and double digestions were both necessary during this investigation. If the enzymes were compatible in the same buffer, a double digestion could be performed. However, if the enzymes required different buffers or conditions, two sequential digestions were performed with a PCR clean-up reaction (Qiagen) in between.

2.6.8 Clean-up of DNA

Between each molecular biology step (e.g. PCR reaction, digestion & ligation), the Qiagen PCR-clean-up kit was used according to manufacturer's instructions to remove any small DNA fragments, impurities and traces of previous buffers.

2.6.9 Ligation of DNA

Ligation of DNA into appropriate vector was performed using a Promega System I kit (according to the manufacturer's instructions).

2.6.10 Transformation

E. coli: Cells were transformed using heat-shock at 42°C with the appropriate plasmid (between 10 - 100 ng), according to the manufacturer's instructions, and grown overnight at 37°C LB with the appropriate antibiotic, Xgal and IPTG combination. Typically 100 $\mu g.ml^{-1}$ ampicillin, 200 $\mu g.ml^{-1}$ Xgal and 0.4 mM IPTG.

Successful transformants were obtained by screening using the selection agar. Several colonies were chosen and colony PCR was performed using the gene specific primers and plasmid primer. This was followed by gel electrophoresis to visualise the result. The plasmid was then harvested using an appropriate plasmid prep method.

C. jejuni: A lawn of the appropriate strain of *C. jejuni* was grown for 24 hours on Brucella medium. The cells were recovered with 1 ml Brucella liquid medium and washed twice campylobacter transformation buffer (CTB) (1 ml, 272 mM sucrose, 15 $ml.100ml^{-1}$ glycerol, 20,000 x *g*, 1 minute, room temperature). In between washes, the cells remained on ice. The cells were resuspended in 100 µl CTB and 10 µl (10 – 100 ng) appropriate plasmid DNA (from Sigma midi-prep) was added to the cells in a electroporation cuvette (Bio-Rad, 2 mm, 0.2 cm electrode). The cuvette was pulsed with 2.5 kV, 200 Ω, load resistance 25 µF Capacitance) using the Gene Pulser 11 electroporation device (Bio-Rad), then 250 µl Brucella liquid medium was added to the cells, and the entire amount was plated onto a pre-warmed (37°C) Brucella plate

and incubated (5 hours, micro-aerobic cabinet), then recovered and re-plated onto a Brucella plate containing the appropriate antibiotic. Transformed colonies appeared after a couple of days incubation in the micro-aerobic cabinet.

2.6.11 Plasmid DNA Preparation

Plasmids were harvested from *E. coli* cultures grown overnight in the appropriate antibiotic using either a Qiagen mini-prep, Qiagen mini-spin-prep or Sigma midi-prep kit depending on the quantity of plasmid required. The plasmids were confirmed as being the correct construct by either digestion followed by electrophoretic separation and sequencing using gene and plasmid specific primers.

2.6.12 DNA Sequence Analysis

Sequencing reactions were either performed by the staff at TGAC (The Genome Analysis Centre) or using the BigDye 3.1 kit (Applied Biosystems) and the reaction was performed according to the manufacturer's instructions. The primers used were appropriate to the gene or plasmid being investigated. Firstly, a 1 minute initial denaturation step at 96°C was followed by 25 cycles of 96°C for 30 seconds, 50°C for 15 seconds and 60°C for 4 minutes. The reaction was held at 4°C.

Constructs were sequenced at TGAC using the AbiPrism 3730 high throughput capillary sequencer. The sequences were interpreted with Vector NTI analysis by comparison to one another, the published gene and plasmid sequences.

2.7 Identification & Cloning of Homospermidine Synthase (HSS) Genes

The genome of *B. japonicum* had previously been sequenced and a gene had been identified as being similar to other known HSS genes. The gene, NP 774402, was identified using BLAST: Basic Local Alignment Search Tool, (www.ncbi.nlm.nih.gov/blast/Blast.cgi)) on the NCBI database (www.ncbi.nlm.nih.gov).

Bradyrhizobium japonicum (NP 774402): this gene was amplified from genomic DNA by PCR using Takara Ex-Taq (Biowhittaker) and the following primers Forward: 5'-AT**GGATCC**ATGAGCCCCGCCTCG-3' and Reverse: 5'-CC**AAGCTT**TCAACGCACCAGAATATTCCGGA-3' (incorporating restriction sites *Bam*HI and *Hind*III, at the 5' and 3' ends of the gene, respectively). The PCR fragment was ligated into the pGEM-T-Easy vector (Promega). *E. coli* XL2Blue (Stratagene) cells were transformed with the plasmid (according to the manufacturer's instructions). The constructs were verified by sequencing (BigDye 3.1, Applied Biosystems) using the ABI3730 high-throughput sequencer, before being digested with *Bam*HI and *Hind*III and ligated into the corresponding sites of the protein expression vector pET21a (Novagen).

Paramecium tetraurelia [XP_001452042], Ralstonia phage RSL1 [YP_001949984], Sawyeria marylandensis, Opitutus terrae PB90-1 [YP_001819503], GOS Marine Metagenome [ECW30472]: The coding region of these 5 genes were synthesised *de novo* with codon usage optimised for protein expression in *E. coli* and *Saccharomyces cerevisiae* (Genscript Corporation, Piscatway, NJ, USA; www.genscript.com.) In each case, *Bam*HI and *Xho*I sites were incorporated at the 5' and 3'end, respectively. The synthetic genes were ligated as *Bam*HI/*Xho*I fragments into the similarly digested pET21a expression vector (Novagen). The *P. tetraurelia* gene was cloned into pET21a whereas the other four were cloned into pET21b to allow for different reading frames of the synthesised genes.

All six constructs expressed the N-terminal T7 tag present in pET21, but contained a stop codon at the 3' end of the HSS gene, and therefore did not contain the C-terminal His-Tag in the expressed protein. All constructs were verified by sequencing prior to protein expression studies.

2.8 Cloning of Cj1501

Campylobacter jejuni NCTC 11168 gene Cj1501 was amplified from genomic DNA by PCR using Phusion Taq (New England Biolabs) and the following primers Forward: 5'- AT**GGATCC**ATCGTGAAAATTACTTA-3' and Reverse: 5'-

AT**CTCGAG**TTAAGGCTTTTGGATTAA-3' (incorporating restriction sites *Bam*HI and *Xho*I, at the 5' and 3' ends of the gene, respectively). The PCR fragment was ligated into the pGEM-T-Easy vector (Promega). *E. coli* XL2Blue (Stratagene) cells were transformed with the plasmid (according to the manufacturer's instructions). The constructs were verified by sequencing (BigDye 3.1, Applied Biosystems) using the ABI3730 high throughput sequencer, before being digested with *Bam*HI and *Xho*I, and ligated into the corresponding sites of the protein expression vector pET21a (Novagen).

The Cj1501 expressed the N-terminal T7 tag present in pET21a, but contained a stop codon at the 3' end of the gene, and therefore did not contain the C-terminal His-Tag in the expressed protein.

2.9 Mutant making in *C. jejuni*

To make mutants in *C. jejuni* NCTC 11168 primers were designed for the target gene (Gene) to amplify both the beginning and end flanks of the gene. The primers contained a linker region (5'-GACGC**GGATCC**GCGTTGGCGCGCCA-3'). A two-step PCR reaction allowed the two flanks to be amplified, then fused together using the linker region. This new fragment was then digested with *Bam*HI and a *Bam*HI-digested kanamycin resistence gene (Kan) was ligated into the fragment. The complete Gene-Kan fragment was ligated into pGEM-T-easy and transformed into *E. coli* Top10 cells (Novagen), where it was amplified and checked by sequencing to contain the desired fragment. The pGEM-Gene-Kan plasmid was then transformed into wild-type *C. jejuni* and by homologous recombination, the Kan was inserted into the genome in the target gene, thus disrupting the gene. The mutants were screened by antibiotic resistance and confirmed by PCR.

2.10 Complementing mutants and wild-type *C. jejuni*

To genetically complement the *C. jejuni* mutants or wild-type strain, specific primers for the desired genes were designed and used to amplify the target gene, or section of genome, with appropriate restriction sites. The gene, or section of genome, was amplified, digested and ligated into plasmids designed

for this purpose which contained a chloramphenicol resistance gene, part of the pseudogene Cj0046 and a *C. jejuni* native promoter, such as *fdx*A or *met*K. The plasmids were transformed into *E. coli* Top10 cells (Novagen), where it was amplified and checked by sequencing to contain the desired fragment. The plasmids were then transformed into the appropriate *C. jejuni* mutant or wild-type strain. The plasmid intergrated by homologous recombination into the pseudogene (Cj0046) area of the genome, and the *C. jejuni* was able to express the new gene with no disruption to the rest of the genetic material. The complemented cells were screened by antibiotic resistance and confirmed by PCR.

2.11 *R. leguminosarum* Mutant Identification

The *R. leguminosarum* Tn5 insertion library at The John Innes Centre (JIC) contains 48 pools of genomic DNA, these pools were screened using primers which annealed internally to the target gene, HSS (5'-GCAAAGCTGATGAAGAAG-3' and 5'-TTCTTCATCAGCTTTGCC-3'), and a Tn5 primer (5'-GAACGTTACCATGTTAGGAGG-3'). After identifying a single pool as containing a potential mutant, the colonies within the pool were screened by colony PCR. The position of the Tn5 insertion was confirmed by sequencing. The mutant phenotype was examined by light microscopy and inoculated onto several pea plants using a facility at JIC.

2.12 Analytical Biochemistry

2.12.1 Polyamine Extraction from Bacterial Cells

Cell pellets (*E. coli, R. leguminosarum* and *C. jejuni*) were collected by centrifugation of 1 ml liquid culture, and washed with 0.5 vol 20 mM Tris pH 6.8 or 8.0. These cell pellets could be stored long-term at -80°C. The polyamines were extracted using 0.2 vol tricholoracetic acid or 0.2 vol Lysis Buffer (100 mM MOPS, 50 mM NaCl, 20 mM $MgCl_2$) with three freeze-thaw cycles. The suspension was left on ice for five minutes and then centrifuged (18,000 x g, 10 minutes) to remove cell debris. The clear supernatant was analysed for polyamine content.

2.12.2 Labelling Polyamines

In order to detect the polyamines during HPLC analysis a labelling group had to be added to each molecule. Polyamines were either labelled using the AccQ●Fluor™ Reagent kit (Waters) or by a benzylation method (both detailed below). Benzyloation was appropriate for LC-MS/MS, whereas AccQ●Fluor™ gave better resolution between polyamines for HPLC alone.

***AccQ●Fluor™*:** Polyamines were derivertised for HPLC analysis using the AccQ●Fluor™ Reagent kit (Waters) following manufacturer's instructions. In each case 5 µl of cell extract was included in the reaction, along with an internal standard, at a concentration of 1.25 µM. Reactions were performed in AccQ●Fluor Borate Buffer in a total volume of 100 µl, and heated to 55°C for ten minutes. The amine groups became labelled with the reagent 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate (AQC).

***Benzoylation*:** 250 µl of the polyamine containing supernatant was added to 1 ml 2 N NaOH and 10 µl benzoylchloride in long glass tubes. The sample was vortexed for ten seconds and then incubated at room temperature for 20 minutes. Then, 2 ml saturated NaCl was added, vortexed (10 seconds), and 2 ml diethylether was added and vortexed again (2 x 10 seconds). The ether phase (1 ml) was then transferred to a fresh glass tube and evaporated under helium gas. Benzoylated polyamines were re-suspended in 100 µl acetonitrile (ACN). The samples were filtered with a 0.45 µm membrane (Sartorius Stedim Biotech) before injection into the HPLC.

Polyamine standards of putrescine (PUT), diaminopropane (DAP), cadaverine (CAD), spermidine (SPD), spermine (SPM) and homospermidine (HSPD) were also labelled for detection during the analytical procedures.

2.12.3 High Performance Liquid Chromatography (HPLC)

***AccQ●Fluor™ labelled polyamines*:** The derivatized polyamine samples were analysed by HPLC using a reverse-phase C18 column (Phenomenex, Luna 5 micron) on a Dionex Summit System. For each sample, 10 µl of derivatised cell extract was injected by an automated sample injector (Dionex ASI-100) and

the polyamines were separated based on hydrophobicity. The system was operated at 33°C and equilibrated with Eluent A (70 mM acetic acid, 25 mM triethlyamine, pH 4.82) at 1.2 ml.min^{-1}. Elution was performed using the following linear gradients of Eluent B (80% acetonitrile): 22% for 5 minutes, 39% for 12 minutes with 6% methanol, 33% for 30 seconds with 14 % methanol, 10% for 6.5 minutes with 70% methanol and finally 100% for 21 minutes (Dionex P580 Pump). The polyamines were monitored by fluorescence (Dionex RF 2000 detector) with a 248 nm excitation filter and a 398 nm emission filter.

***Benzoylation labelled polyamines*:** The benzoylated polyamine samples were subjected to HPLC (Summit System, Dionex) equipped with an automated sample injector (Dionex ASI-100) and a reverse-phase C18 column (Luna 5 micron). 1 mM diaminoheptane (DAH) was used as an internal standard. The HPLC ran with a acetonitrile:water mixture with a flow rate of 1.2 ml min^{-1} (Dionex P580 Pump). The acetonitrile gradient was 50% for 8 minutes, 80% for 7 minutes and 100 % for 15 minutes. Polyamines were detected using a UV detector at 254 nm (Dionex RF 2000 detector).

2.12.4 Identification of Polyamines by HPLC

Polyamines were identified by comparison of retention times with known standards which were derivatised and run on the HPLC alongside cell extracts. Precise quantification of HSPD was performed using the internal standard and a calibration curve of HSPD (0.05 μM to 50 μM). Polyamines were quantified by analysing the peak area with respect to the internal standard, then adjusted by the cell pellet weight. Three technical repeats were used for quantification.

Further confirmation of the HPLC results and the identification of unknown polyamines were performed by LC-MS/MS.

2.12.5 Liquid Chromatography - Tandem Mass Spectroscopy (LC-MS/MS)

Samples were benzoylated as described above then analysed using an Agilent 1200 Binary Liquid Chromatograph coupled to an Applied Biosystems 4000

QTRAP triple quadrupole mass spectrometer (Warrington), equipped with electrospray ionisation (ESI) source operated in positive ion Enhanced Mass Spectrometry (EMS) mode and diode array detector collecting 190-400 nm.

Chromatography was performed (as described above) on a reversed phase Luna 5 micron C18(2) column, 100 Å, 250 x 4.6 mm (Phenonenex, UK) at 30°C. The mobile phase was MilliQ water and acetonitrile at a flow rate of 1.2 ml.minute^{-1} and split 50:50 to the diode array and mass spectrometer. The gradient started with 50% B, changing to 80% over eight minutes and then to 100% B at 15 minutes and held until 30 minutes. Injection of 25 µL was made following a ten minute equilibration between injection cycles.

For the mass spectrometer, the TurboIonSpray voltage was set at 4,500 V, temperature was at 550°C, curtain gas was at 1 KPa nebulizing gas pressure was at 3 KPa, drying gas pressure was at 2 KPa, declustering potential was at +75 V, entrance potential was at +10 V, collision gas was at a high flow, and collision energy was at 10 eV. The ion trap was operated in dynamic fill time mode. Enhanced mass scan data were acquired using the Q3 as a linear ion trap for 35 minutes over a scan range of m/z 50–2000 with a scan rate of 1000 Da.s^{-1} and a step size of 0.08 Da.

2.12.6 Nuclear Magnetic Resonance Spectroscopy (NMR)

NMR was used to identify the presence, absence and concentration of several metabolomic compounds in *C. jejuni* growth medium. The medium was filter-sterilised (0.2 µm) to remove the cells, then diluted 0.5 fold with a D_2O containing-buffer. High resolution ^{1}H NMR spectra were recorded on a 600 MHz Avance spectrometer (Bruker) fitted with a cryoprobe and a 60 slot auto-sampler. Each spectrum consisted of 128 scans of 32768 complex data points with a spectral width of 14ppm. The *noesypr1d* pre-saturation sequence (Bruker) was used to suppress the residual water signal with low power selective irradiation at the water frequency during the recycle delay and mixing time. Temperature: 27°C; Internal lock: D_2O; Scan number: 128 scans; Data points per spectrum: 32768; Spectral width: 8389.262 Hz; Pulse time: 8.8 µseconds (90^0); Acquisition time: 1.5 sec; Mixing time: 0.15 sec; Recycle delay:

2 sec. Spectra were transformed with 1Hz line broadening, manually phased and baseline corrected using the TOPSPIN 2.0 software.

2.13 Protein Chemistry

2.13.1 Induction of Protein

Cultures of *E. coli* BL21(DE3) (Novagen) containing the pET21a plasmid containing the gene of interest or an empty vector were grown in triplicate cultures for each construct in an appropriate volume of LB broth containing 100 $\mu g.ml^{-1}$ ampicillin. Protein expression was induced using Novagen Overnight Express™ Autoinduction System 1 and growth at 25°C for 24 hr. Cells were harvested in appropriate volume aliquots by centrifugation, washed twice in 0.5 vol PBS (8 g NaCl, 0.2 g KCl, 1.44 g Na_2HPO_4, 0.24 g KH_2PO_4 in 1 L water). The cell pellets were stored at -80°C.

2.13.2 Protein Extraction

Either sonication or a Bugbuster (Novagen) method was used to liberate the soluble protein from the bacterial cells.

Sonication (3 x 15 seconds) (Sanyo Soniprep 150 MSE) of cell samples (*C. jejuni, R. leguminosarum* and *E. coli* vol 50 - 1000 µl) with incubation on ice between sonications followed by gentle centrifugation (6000 x *g*, 30 seconds) to remove the cell debris removed the cellular protein for visualisation on a gel and for the FDH assay.

Bugbuster (Novagen) uses the enzymes benzonase, nuclease and rLysozyme to gently disrupt the cell wall of *E. coli* to liberate soluble proteins. The method was followed according to the manufacturer's instructions.

2.13.3 Protein Quantification

To gain a rough estimate of protein concentration, soluble protein and purified recombinant protein were added to Nu Page x 4 loading buffer (Invitrogen), denatured by heat treatment (70°C for 3 minutes) and run on a

10% SDS Nu-Page Bis-Acrylamide Gel (Invitrogen) alongside a broad range molecular weight marker (Promega).

The Bradford method of protein quantification was used. This involved using 1 in 5 diluted Bradford Reagent (Bio-Rad) and bovine serum albumin (BSA) as a standard. BSA (at a range of known concentrations 0.2 - 1.0 $mg.ml^{-1}$) and the experimental protein was incubated in the reagent for 10 minutes and the absorbance at 595nm correlated to the concentration of protein.

Before two dimensional (2D) protein separation, another method (2D Quant Kit, Amersham) was used to gain a more accurate protein concentration (see section 2.14.2). The reagents in this method were compatible with the rest of the 2D proteomic method.

2.13.4 Protein Purification

The pET21 vector contains a T7 affinity tag of eleven amino acids which is fused to the N-terminus of the protein as it is synthesised. Using the Novagen system, the target protein can be purified. Purification of the tagged protein is based on binding target proteins to the T7-tag monoclonal antibody which is covalently coupled to cross-linked agarose beads. Unbound proteins were washed away initially and the tagged proteins were eluted with a separate buffer. The procedure was followed according to the manufacturer's instructions.

2.13.5 Protein Buffer Exchange

A 20 mM Tris pH 8 buffer containing 20% glycerol was used to store the protein at -80°C. The protein was transferred to the buffer by elution from a Sephadex PD-10 column.

2.13.6 Protein Concentration

A centrifugal Millipore filter device (AmiconUltra – Ultracel 10K) filtered and concentrated the HSS protein (50 kDa) solution. A smaller, Biomax-5K NMWL membrane filter (Millipore) was used to concentrate the Cj1501 protein (8 KDa).

2.14 Proteomics

2.14.1 Cell Growth and Protein Extraction

C. jejuni cells were harvested from 50 ml Brucella broth by centrifugation (4,000 x *g*, 10 minutes, room temperature). The cells were frozen immediately and care was taken to ensure little technical difference between samples.

The cell pellets were resuspended in 500 µl IPG Lysis Buffer (7 M Urea, 2 M Thiourea, 4% CHAPS, 40 mM Tris Base) and sonicated (Soniprep 150 MSE, Sanyo) for 10 seconds six times with incubation on ice between sonications. 20 minutes incubation on ice following sonication. The samples were centrifuged (20,000 x *g*, 20 minutes, 4°C) and aliquoted (50 µl) for long term storage at -80°C.

2.14.2 Protein Quantification

Protein concentrations for 2D gel samples were determined using the 2D Quant Kit (Amersham) according to the manufacturer's instructions. An appropriate amount of sample (5 and 10 µl) and BSA, used as standard curve (5 vol, 0–25 µl, 2 $mg.ml^{-1}$), was added to 500 µl precipitant (provided in 2D Quant Kit), vortexed and incubated (3 minutes, room temp.). 500 µl co-precipitant was added to every sample and standard, vortexed (10 seconds), centrifuged (20,000 x *g*, 5 minutes, 4°C). The pellet was resuspended into 100 µl copper solution and 400 µl water. Then, 1 ml working colour reagent (100 part solution A, 1 part solution B from kit), was added to each sample and inverted immediately. After 15-20 minutes the absorbance was measured at 480 nm (Spectrophotometer Uvikon XL, NorthStar Scientific) and a standard curve was general using the know BSA standards. The assay was performed in duplicate for each sample.

2.14.3 Protein Concentration and Dialysis

Protein was concentrated to 50 µl or less using a Biomax-5K NMWL membrane filter (Millipore) to gain a volume of protein solution containing 100 µg protein (10,000 x *g*, 4°C). 450 µl non-bromophenol blue rehydration

buffer was added and the sample mixed, then centrifuged using the same membrane filter (10,000 x *g*, 20°C) until the volume ~50 µl. Then, 400 µl rehydration buffer was added to the sample ensuring all the protein was collected from the membrane filter.

2.14.4 SDS-page Polyacrylamide Gel Formation for 2D Gel Electrophoresis

To make eleven 10% gels: 344.75 ml 30% Duracryl 0.65% Bis (Genomic Solutions), 257.79 ml 1.5 M Tris slab gel buffer and 425.50 ml water were mixed and de-gassed for 10 minutes. Sequentially SDS solution (10% 10.87 ml), TEMED (0.520 ml) and 10 % Ammonium Persulphate (2.64 ml) were added and mixed to the solution.

The 2D 10% Duracryl gels SDS-page gels (28 x 23 cm; 1 mm thick) were made by pouring the SDS-page acrylamide solution between two custom-made glass plates in a specially designed chamber (Investigator, Genomic Solutions). Each gel was overlaid with 1 ml water saturated butanol and the gels polymerised after 3-4 hr. Each gel was then cleaned to remove the butanol and the excess polyacrylamide gel. The gels were stored in sealed bags with ~50 ml 1/5 diluted Tris slab gel buffer for up to 2 weeks at 4°C.

2.14.5 1st Dimension Electrophoresis

100 µg protein was made to volume 400 µl with IPG strip rehydration buffer (7 M Urea, 2 M Thiourea, 2% CHAPS, 0.1% Bromophenol Blue, 28 $mg.ml^{-1}$ DTT) and loaded into reswelling tray. The 18 cm 3–10NL Immobiline DryStrips (Amersham Biosciences) was laid into the solution. The gel strip was rehydrated at 20°C for minimum 12 hr. A moist electrode wick was placed at either end of the IPG gel strip and each well was covered with mineral oil ~2.5 ml. Isoelectrical focussing was performed for 80 kVh at 20°C over 24 hours using the pHaser system (Genomic Solutions). 1st dimension gel strip can be stored at -80°C.

2.14.6 2nd Dimension Electrophoresis

Prior to the second dimension the focused strips were incubated in filtered (0.45 µm, Sartorius Stedim Biotech) equilibration buffer (5% SDS and 0.01% bromophenol blue in 0.122 M Tris/acetate, Genomic Solutions). To reduce and alkylate cysteines the strips were treated first with 8 $mg.ml^{-1}$ DTT in the equilibration buffer (9 ml; 30 minutes with gentle rotation [Rotatest shaker, R100/TW, Luckham]). Then the strips were transferred into 25 $mg.ml^{-1}$ iodoacetamide containing equilibration buffer (9 ml; 30 minutes with gentle rotation). The second dimension separated the proteins based on molecular weight (kDa). The Investigator 2nd Dimension Running System (Genomic Solutions) was used with cathode buffer (200 mM Tris base, 200 mM Tricine, 14 mM SDS,) and anode buffer (25 mM Tris-acetate buffer, pH 8.3). Electrophoresis was carried out using either a maximum voltage of 500 V or a maximum power of 20 W per gel.

2.14.7 Imaging 2D Gels

The gels were carefully removed from the glass plates. The proteins were fixed (400 ml, 40% methanol, 10% acetic acid, overnight). Then proteins were stained by Sypro-Ruby (Bio-Rad, 400ml, overnight, in the dark) according to the manufacturer's instructions. Finally, the gels were de-stained (400 ml, 10% methanol, 6% acetic acid, 1 hour minimum). The gels were imaged at 100 µm resolution using the Pharos FX Laser Scanning Fluorescent Imager (Bio-Rad) and a 16-bit image in 65,000 shades of gray was obtained. The laser strength (voltage, % of maximum strength) for the Photo Multiplier Tube (PMT) was manipulated for each image to improve image quality. The excitation filter used was 532 nm. Gel images were compared using ProteomWeaver analysis software (Definiens). The gels were stored in the de-stain, in the dark at 4°C, for several weeks.

2.14.8 Spot Pick and Protein Identification

Proteins of interest were removed from the gel using ProPick excision robot (Genomic Solutions), and in-gel trypsin digested using a ProGest Protein Digester (Genomic Solutions). Gel plugs were conditioned with two 20 minutes incubations in 200 mM ammonium bicarbonate (ABC) in 50%

acetonitrile (50 µl), followed by 10 minutes incubation with acetonitrile (50 µl). The gel plugs were then conditioned for 15 minutes with 25 mM ABC (50 µl) followed by 10 minutes in acetonitrile (50 µl). A final 5 minutes incubation of acetonitrile (50 µl) preceded the trypsin digestion at 37°C (3 hr) using 50 ng (5 µl per well) sequencing-grade porcine trypsin (Promega) dissolved in 25 mM ABC. Digestion was stopped and peptides extracted using formic acid (5%; 5 µl per tube). The tryptic digests were analysed using a Reflex III MALDI-TOF instrument (Bruker) with a Scout 384 ion source using a nitrogen laser (wavelength 337 nm) to desorb/ionize the matrix/analyte material from the sample substrate. The matrix used was a mixture of a-cyano-4-hydroxycinnamic acid and nitrocellulose, applied to the target as a thin film. Sample (0.3 µl) was then spotted on to the target and dried before washing with 10% formic acid. All spectra were acquired in a positive-ion reflector. The acceleration voltage was set to 25 kV, the reflection voltage to 28.7 kV, the ion source acceleration voltage to 20.9 kV, and the reflector-detector voltage to 1.65 kV. Proteins were identified from the peptide mass peak list by the Protein Mass Fingerprint technique using the Mascot search tool (Matrix Science; www.matrixscience.com/). *C. jejuni* protein sequences in the National Center for Biotechnology Information (NCBI) database (www.ncbi.nlm.nih.gov/).

2.15 Enzyme Assays

2.15.1 Homospermidine Synthase (HSS) Activity

HSS activity was measured by determining the amount of homospermidine formed (detected by HPLC analysis). The assay had a volume of 100 µl containing: 0.5 mM NAD^+, 5 mM Putrsecine, 70 mM KCl, 1 mM DTT, 50 mM Tris pH 8.6 and approximately 20 µg enzyme. The reaction was started by the addition of the enzyme and incubated at 28°C over a time course. A control sample was also taken initially (Time = 0 hour). To characterise the enzyme; pH, temperature and substrates were altered. The reaction was stopped by addition of 1 ml 2N NaOH. A blank reaction was also run without enzyme.

2.15.2 Ammonia (NH_3) Production

An Ammonia Assay Kit (Sigma) was used according to the manufacturer's instructions to test the production of ammonia (NH_3) in recombinantly purified protein samples. The reaction was mixed as in the HSS assay. 50 µl of this mixture was added to a cuvette containing the ammonia assay reagent. A blank of water and an ammonia standard were also used. The samples were incubated at room temperature for 5 minutes and 10 µl L-Glutamate Dehydrogenase was added, mixed and incubated at room temperature for 5 minutes. The absorbance at 340 nm was measured using Spectrophotometer Uvikon XL (NorthStar Scientific, UK). Using an equation provided by the manufacturer, the concentration ($\mu g.ml^{-1}$) of NH_3 produced during the HSS assay was calculated.

2.15.3 Ornithine Decarboxylase (ODC) Activity

Cells pellets of *R. leguminosarum,* grown in 10 ml AMS liquid medium, were resuspended in 500 µl ODC Buffer (0.1 M HEPES, 10 mM EDTA, 0.5% ascorbic acid, pH 7.4, 10mM DTT, 1 x pyridoxal 5-phosphate, 5% polyvinyl polypyrrolidone) and kept on ice until use. Radioactive ornithine substrate was prepared (0.1 µCi L-[1-^{14}C]ornithine (57 $mCi.mmol^{-1}$) and 2.5 mM unlabelled ornithine) and stored on ice until use. 300 µl 1 M KOH was added to a scintillation vial along with 20 µl substrate. 100 µl sample (in an open eppendorf) was placed in the scintillation vial and the reaction progressed in the sealed vial at 37°C for 45 minutes. The reaction was stopped by the addition of 250 µl 5% TCA directly into the eppendorf using a syringe. The incubation continued at 37°C for a further 45 minutes. The eppendorf tube was removed and 5 ml scintillation fluid was added to the vial, vortexed to mix and left for 1 hour. The liquid scintillation counter (Tri-carb 2700 TR, Packard) measured the release of CO_2 from L-[1-^{14}C]ornithine and the ODC activity ($nmolCO_2.hr^{-1}.mg^{-1}$) in each sample was determined. Protein concentration was determined using the Bradford assay (as described previously).

2.15.4 Formate Dehydrogenase (FDH) Activity

The FDH activity assay was adapted from a method supplied by Dr. Dave Kelly, Sheffield. Bacterial cells were grown over-night in Brucella broth and the OD_{600nm} was measured. 5 ml of culture was spun down (10 minutes, 10,000 x *g*) and washed twice with 1 ml 25 mM phosphate buffer pH 7.4 (centrifuged for 1 minute, 10,000 x *g*). The cells remained on ice until the start of the enzyme assay. The cells were resuspended in 1 ml 25 mM phosphate buffer pH 7.4 and 200 µl was aliquoted into the anaerobic cuvettes. Into the cuvettes went 200 µl 10 mM benzylviologen and 1600 µl buffer. The cuvettes were individually sparged with nitrogen for 8 minutes. The cuvette was then heated to 37°C and a reading was taken at 578nm. The substrate was then added; 20 µl 1 M Sodium formate sparged with nitrogen for approx 20 minutes prior to use. The reaction was monitored at 578 nm at one reading per second for approximately 200 sec, or until the reading was above 2.0.

2.16 Transcriptional Analysis by Microarray

The microarray slide was purchased to order from Agilent, USA (code: G2509F). The slide is specific for the genome of *C. jejuni* NCTC 11168 and was designed by Bruce Pearson, IFR. Each gene had four to five probes (depending on gene length) in duplicate and when the RNA specific to that probe was present, it bound during the hybridisation step. Complementary DNA was produced from the RNA using the random hexamer provided in the Affinity Script kit (GE Healthcare). To visualise the DNA during the microarray, it was labelled with single-stranded DNA containing the Cy3 or Cy5 flurescence marker. The Affinity Script was mixed with the RT buffer, 0.1 M DTT and dNTPs according to the manufacturer's instructions and incubated for 10 minutes at room temperature, then 42°C for 16 hr.

Each sample was then subjected to a Qiagen PCR purification method, using the provided columns and according to the manufacturer's instructions. The samples were then dehydrated using a Savant Speedivac Plus SC2104 (Thermoquest) with attached Vacuum System UV5400A (Thermoquest).

The dried samples were resuspended in 10 µl diethylpyrocarbonate (DEPC) treated water and 25 µl 2 x HiRPM Hybridisation buffer (Aglilent) and 5 µl blocking agent were added (Agilent Hi-RPM Gene Expression Hybridisation Kit). The samples were heated to 103°C (3 minutes) and 40 µl was added onto the array. The array was securely fastened into the gasket and hybridisation occurred overnight in the hybridization oven chamber (Shel Lab) with constant rotation.

The array slide was removed from the chamber constantly submerged in Agilent gene expression wash buffer 1, incubated for 5 minutes in the buffer 1 on a roller mixer SRTI (Stuart) (room temp., in dark), incubated for 2 minutes in Agilent gene expression wash buffer 2 (37°C, in dark). Followed by rinses in acetonitrile and stabilisation solution 1 sequentially.

DNA microarrays were scanned using an Axon GenePix 4000A microarray laser scanner (Axon Instruments) and the data from detected features initially processed using the GenePix 3.0 software.

2.17 Bioinformatics

The bioinformatic techniques used in Chapter Nine are mainly referenced in text when applicable.

All software packages used in this investigation were freeware available online and a list of the programs used is detailed below:

To search sequenced genomes and protein databases for sequence homology:

BLAST (Basic Local Alignment Search Tool)

http://blast.ncbi.nlm.nih.gov/Blast.cgi

To look at relationships between genes:

String - http://string-db.org/

To determine the domains present in a protein sequence:

Pfam - http://pfam.sanger.ac.uk/

To determine the transmembrane domains in a protein sequence:

TMHMM-2.0 - http://www.cbs.dtu.dk/services/TMHMM/

To look at the annotation and arrangement of the *C. jejuni* and *R. leguminosarum* genomes:

CampyDB - http://www.xbase.ac.uk/campydb/

RhizoDB - http://www.xbase.ac.uk/rhizodb/

To compare two protein sequences for similarity:

SIM - http://expasy.org/tools/sim-prot.html

To compare and align two or more protein sequences:

Tcoffee - http://www.ebi.ac.uk/Tools/msa/tcoffee/

To convert DNA sequence to protein sequence:

Sequence analysis tools - http://expasy.org/tools/

To scan protein sequences for possible motifs:

MyHits motif scan - http://myhits.isb-sib.ch/cgi-bin/motif_scan

Chapter Three

Introduction to Polyamines

TABLE OF TABLES

TABLE OF FIGURES

Chapter Three - Introduction to Polyamines

3.1 Introduction

A polyamine is a naturally occurring hydrocarbon molecule which contains two or more amine groups (NH_2). In living cells, polyamines are found as diamines, triamines, tetramines, pentamines and hexamines, which contain two, three, four, five and six amine groups respectively. Polyamines are polycationic compounds which are present in all biological materials and have been implicated in a wide variety of biological reactions, including the synthesis of deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and protein [1]. Their physiological and biochemical significance in the process of cell growth and differentiation has been repeatedly and extensively reviewed [1-6]. An overview of global polyamine biosynthesis is depicted in Figure 3.1.

During normal metabolism, various cellular macromolecules interact with polyamines electrostatically or covalently. The complexity of polyamine metabolism and the multitude of compensatory mechanisms which are invoked to maintain polyamine homeostasis suggest that these amines are critical to cell survival [6]. An example of this highly regulated pathway is the polyamine-synthetic enzyme ornithine decarboxylase (ODC); which produces putrescine (a widely found polyamine) from ornithine. ODC is regulated on a variety of levels including the necessity of the co-factor being bound [7]. Regulation occurs on a transcriptional level [6] and activity is also subject to feedback regulation in the presence of high or low concentrations of polyamines [8]. Furthermore, ODC degradation is itself regulated, and is accelerated by increases in the levels of cellular polyamines [8].

Polyamines are found in all living species, except two orders of Archaea, Methanobacteriales and Halobacteriales [9]. This conservation across evolution is a positive feature in that it argues for their importance in cell survival [6], but the variety of polyamine profiles seen (i.e. different polyamines in different species) implies a lack of specificity to polyamine function.

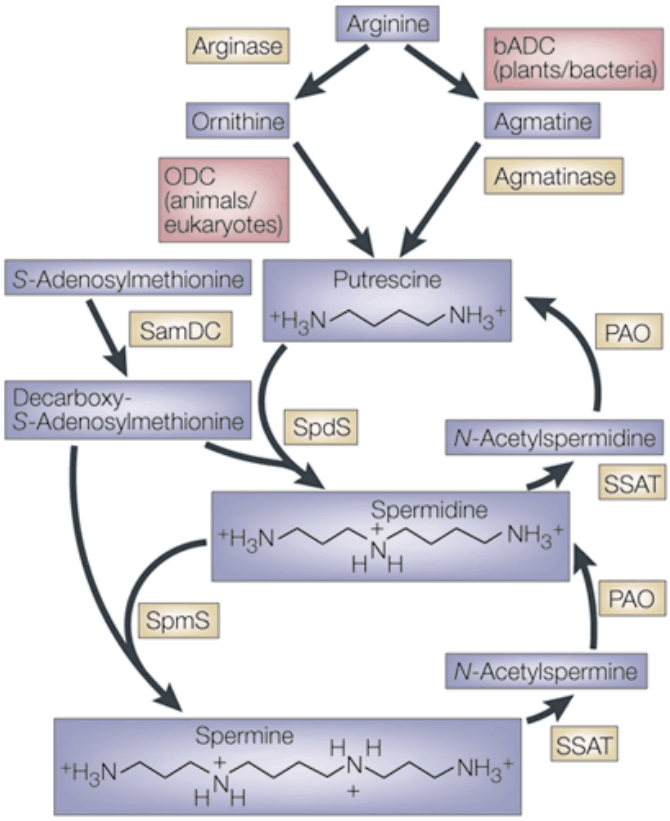


Figure 3.1 An overview of common polyamine biosynthesis.
Ornithine decarboxylase (ODC) initiates the production of polyamines by producing the diamine putrescine. In some plants and bacteria arginine decarboxylase (bADC) initiates an alternative two-step pathway to putrescine. Next, an aminopropyl group derived from methionine by the decarboxylation of *S*-adenosylmethionine adds to putrescine to form spermidine and a second aminopropyl group adds to spermidine to form spermine by discrete synthases, spermidine synthase (SpdS) and spermine synthase (SpmS). The biosynthetic process can be reversed by a series of catabolic reactions that are initiated by spermidine/spermine acetylase (SSAT), a step that is rate-limiting for conversion of spermine to spermidine and of spermidine to putrescine. Acetylation is followed by the action of polyamine oxidase (PAO). The three key enzymes that control the rate of flux through this pathway, ODC, *S*-adenosylmethionine decarboxylase (Sam-DC) and SSAT, are each subject to multilevel control, with degradation being especially prominent. Entry to and exit from the cell also affect polyamine pools. Taken from [10]. Note: homospermidine is not shown on this diagram as it is classed as an 'uncommon' polyamine.

3.2 Common & Uncommon Polyamines

Putrescine, spermidine and spermine occur ubiquitously in various combinations in higher living organisms; these are referred to as the 'common' polyamines. Most prokaryotic and eukaryotic organisms synthesise putrescine and spermidine. Figure 3.2 summarises some of the linear

polyamines discussed in this chapter. The diversity in carbon chain length and amine-group position can clearly be seen.

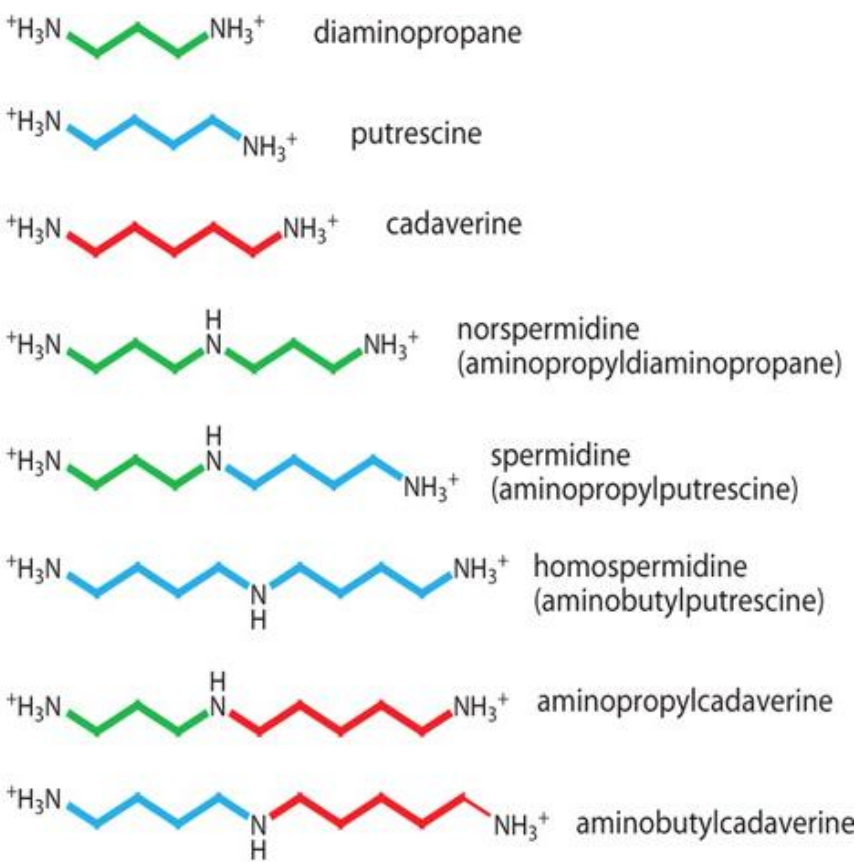


Figure 3.2 A skeleto-chemical representation of common and uncommon polyamines.
The colour of the carbon chain represents the number of carbons in a continuous chain between nitrogen atoms. Green, 3; Blue, 4; Red, 5.

Certain polyamines in plants and micro-organisms which are structurally different to the common polyamines, and those generally discovered more recently, are designated 'uncommon' polyamines. The first uncommon polyamines, thermine and thermospermidine, were originally discovered in thermophillic bacteria [11]. Subsequently, uncommon polyamines have been identified in many different types of bacteria, algae, plants, mosses and even viruses [12-14].

Norspermidine, homospermidine, aminopropylcadaverine and thermospermine are just a few examples of uncommon polyamines [15]. Novel tertiary branched tetra-amines, quaternary branched penta-amines, linear penta-amines and linear hexa-amines have been identified distributed

as the major polyamines in six obligately extremely thermophilic eubacteria belonging to Thermoleophilum, Bacillus or Hydrogenobacter [16]. Polyamines, common and uncommon, are distributed throughout the kingdoms of life; some have very simple structures and others are much more complicated. Many of these uncommon polyamines lack the diamino-butylgroup, which is characteristic of the common polyamines [15].

Currently, it is believed that uncommon polyamines have a limited distribution in nature, but this may be untrue and based upon our lack of understanding of polyamine prevalence across all species. As science progresses, our understanding of unculture able bacteria due to metagenomic sequencing improves and we may, in future, discover that these 'uncommon' polyamines are actually the most abundant on the planet. In addition, polyamine detection methods are more sensitive than ever before. Research into polyamine biology is a fascinating field, and the discoveries made during this investigation and others of a similar nature increase our understanding of how and why polyamines function. The wider implications of polyamine research include plant and mammal drug development [17-19].

3.3 Polyamine Function & Fundamental Requirement

The absolute need for polyamines has been identified across several species. For example, polyamine biosynthesis is essential in the fungi *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Aspergillus nidulans* and *Ustilago maydis* [20-23], the parasites *Trypanosoma brucei, Leishmania donovani* and *Giardia lamblia* [24-26]. Polyamines are essential for embryo development in mammals [27, 28] and seed development in plants [29].

The precise function of polyamines within the cell is unknown, however, polyamines have been linked to several different growth and development roles. Mutant *Escherichia coli* strains that do not contain any polyamines can still grow indefinitely, although the rate is only about 30% of the rate observed after exogenous polyamine addition [30, 31].

The uncommon polyamines, some of which have been identified in extreme environments, are believed to have adapted in defined and specialist ways to

serve specific protective roles in both bacteria and plants adapted to extreme environments [12, 32]. An increase in uncommon polyamines was observed when plants were subjected to heat stress [33]. In thermophillic bacteria, higher temperatures lead to the production of polyamines with a larger molecular weight [11, 12].

3.4 Polyamines & Chemical Interactions

At physiological pH, polyamines are positively charged. They carry one positive charge on each nitrogen atom they possess. Understandably, it has been suggested that polyamines are 'supercations' which are equivalent to two or three positively charged ions, such as calcium, or magnesium. However, in polyamines the positive charge is distributed along the entire length of the molecule; this distinguishes them from bivalent cations.

Interactions between cationic polyamines and negatively charged nucleic acids play a pivotal role in DNA stabilisation and RNA processing that may affect gene expression, translation and protein activity [34]. These interactions are affected kinetically by small changes in the microenvironment, such as temperature or pH [15]. The charge on polyamines enables them to interact electrostatically with polyanionic macromolecules, such as DNA, but if charge was the only defining feature of the polyamine, it would not be necessary to have the abundance and differentiation of polyamines which are distributed throughout living organisms. The complex metabolism and regulation of polyamines suggests they have critical and specific functions within the cell which far exceeds the need for a simple charge-charge interaction. Atomic force microscopy has shown that putrescine, spermidine, and spermine self-assemble with phosphate ions to generate nuclear aggregates of polyamines, which may interact with DNA in the cell nucleus [35].

This investigation will explore the need for polyamines with a specific size, shape and charge by choosing model organisms from which the polyamines can be removed from and replaced with an alternative polyamine. The phenotype of these micro-organisms will then be examined.

3.5 Polyamines & The Cell Membrane

Polyamines perform a number of roles by interacting with the cell membrane, as they are able to interact with the phospholipids in the cell membrane [36]. Spermidine and spermine increase the rigidity of the membrane by binding to phospholipids and protein complexes. Polyamines can help prevent the loss of essential nutrients, and in bacterial systems tetramines protect membranes against osmotic lysis [15]. However, in some situations polyamines interfere with the phospholipid membrane and normal permeability of the cell. For example, cadaverine might be a natural regulator of porin activity [37] and polyamines bind to the negatively charged phospholipid head groups or other anionic sites on membranes, thus altering the stability and permeability of such membranes [38]. These interactions can be inhibiting or toxic to the cell. In addition, polyamines are anti-oxidants and prevent lipid peroxidation [39, 40]. Polyamine depletion has also been linked to a reduced phospholipid content [41, 42].

Polyamines affect the regulation of several membrane bound proteins, such as ion channels. The presence of polyamines can be beneficial or detrimental to the cell membrane depending on concentration. For example, spermine (40 mM 1 % [w/v] aqueous solution) can inhibit and dissociate cell-cell membrane interactions in species such as the Gram-negative *Morganella* and *Providencia* [43], and can stabilise membrane-localised enzymes in *E. coli* [41, 44]. This evidence suggests that polyamines have a complex role within the phospholipid membrane, but due to the uniqueness and differentiation between species, it is impossible to apply any general rules to the polyamine and cell membrane relationship.

3.6 Polyamines & Ionic Cellular Concentrations

Polyamines are involved in diverse processes, including the regulation of gene expression, translation, cell proliferation, modulation of cell signalling and membrane stabilization [45]. Because of these multifaceted functions, the homeostasis of polyamines is crucial and is ensured through regulation of biosynthesis, catabolism and transport [45]. Both plants and animals produce

large amounts of putrescine in response to hyper-osmotic conditions [41]. In the higher organisms, complex homeostatic methods are in place and polyamines are rapidly formed or excreted to maintain cellular pH [41].

Bacteria are able to change their polyamine content in response to ionic challenge. The roles polyamines play in maintaining ionic homeostasis in bacteria are elaborate and complex. Polyamines have a significant role in the growth of *Acidiphilium facilis* 24R under acidic conditions. A mutant of *A. facilis* prepared by UV-irradiation and nitrosoguanidine treatment was identified as being unable to grow under acidic conditions. The growth of the mutant could be fully restored by the addition of spermidine or lysine at the concentration of 100 μM [46]. The putrescine content is high in organisms grown at low ionic strength and is much less when these cells are placed in media of higher osmolarity [41].

3.7 Polyamines & Bacterial Growth

Polyamine concentration has a defined effect on cell proliferation. The addition of polyamines into a medium can inhibit bacterial growth, e.g. *Staphylococcus aureus* growth was inhibited by half by approximately 90 μg ml^{-1} of spermine and completely by approximately 120 μg ml^{-1} when the spermine was present at the time of inoculation. This was because the incorporation of amino-acids into proteins was inhibited [47]. But in some cases, polyamines can enhance growth, e.g. spermine (100 μg ml^{-1}) acted as nutritional suppliment when exogenously supplied to *Pasteurella tularensis* [48]. Yoshida *et al.* [49] concluded that several regulators of transcription of some growth-related genes are enhanced by polyamines at the level of translation. In these experiments, *E. coli* were grown in the presence (100 μg ml^{-1}) or absence of putrescine. DNA microarray showed that 309 of 2,742 mRNA species were up-regulated by polyamines.

It is not only the presence of these additional polyamines, but the concentrations which affect growth. Often, as the polyamine concentration increases, so will the proliferation of the cells, until a critical limit is reached, and at this point, the cells will not survive as the polyamine concentration is toxic.

It is hypothesed that the intracellular concentration of polyamines may participate in the control of the synthesis of ribosomal RNA in bacteria [50]. The natural polyamine content of cells during the growth stage is proportional to the mass of the cells, and the polyamine content increases as the cells replicate, but remains stationary when cell division ceases [50]. However, cells can be artificially manipulated by the addition of certain amino acids, and polyamine concentration can increase without the change in growth rate. In *E. coli*, the supplementation of methionine (20 μg ml^{-1}) leads to spermidine concentration increases without the corresponding growth rate change [50].

3.8 Polyamines & Cellular Death

Polyamines have historically been associated with cell growth, but more recently, it has been shown that polyamines can also regulate controlled cell death (apoptosis) [51]. Polyamines and many of their structural analogues exert cytotoxic effects *in vitro* as well as *in vivo*, this effect has been observed during cell death of nerve cells, in programmed cell death of embryonic cells and in various *in vitro* models of apoptosis [51]. Polyamines are bivalent regulators and can promote cell growth or death, depending on other cellular, chemical and environmental signals [6].

3.9 Polyamines & DNA

Spermidine and spermine can bind within the groves of DNA and clamp the individual molecule, or two different DNA molecules, together [52, 53]. In the DNA structure, it has been reported that polyamines have an affinity for pyrimidine residues, thymidine and cytosine [6, 41]. Polyamines are able to alter the DNA-nuclear matrix interaction, and it is suggested that polyamines can alter the structure of DNA [54]. The polyamine analogue, 1,14-bis(ethylamino)-5,10-diazatetradecane, causes DNA to negatively supercoil more than the control, and ultimately, this leads to inhibition of the growth of tumor cells in tissue culture [54].

Spermidine in high concentrations (greater than 2 mM) can inhibit DNA repair by DNA polymerase I, but in lower concentrations, both putrescine and spermidine enhanced the activity [41]. Polyamines do not stimulate incision nucleases, which remove and repair UV damaged DNA [55]. DNA ligase and polynucleotide kinase are both increased in activity by low concentrations of polyamines [41].

3.10 Polyamines & Phylogenetic Studies

In bacteria, polyamine composition is unique on the level of order, family or genus. The phylogenetic relationship and evolutionary divergence among bacterial species can be estimated by the polyamines present [56]. The field of comparative genomics is changing our understanding of evolution, phylogeny and microbial pathogenesis [57], and polyamine based phylogenetic studies within species are becoming increasingly common.

Recently, there was a large phylogenetic analysis which featured over 200 eukaryotic homologues of ornithine decarboxylase (ODC) and evaluated their potential to be either true ODCs or catalytically inactive proteins that might be analogues of a previously identified antizyme inhibitor. Antizyme is a protein that inhibits ODC activity and is induced after addition of polyamines to cells. The investigation was a bioinformatic experiment, and they did not test the ODC activity biologically, but merely examined and compared protein sequences [58]. This analysis yielded several orthologous groups of putative novel antizyme inhibitors each apparently arising independently. The investigation also identified previously unrecognised ODC paralogues in several evolutionary branches, including a previously unrecognised ODC paralogue in mammals. However, to validate the conclusions of this investigation, the enzyme activity should be validated by experimentation.

There have been several published phylogenetic studies relating to evolution and occurrence of polyamine synthetic enzymes, such as the study of putrescine-N-methyltransferase and spermidine synthase [59] and S-adenosylmethionine decarboxylase, which is involved in the synthesis of spermidine and spermine [60].

The investigation of Biastoff *et al.* [59] concluded that putrescine-N-methyltransferase evolved from the ubiquitous spermidine synthase. Analysis of the spermidine synthase protein structure suggested that only a few amino acid exchanges in the active site were necessary to achieve putrescine-N-methyltransferase activity. Protein modelling, mutagenesis and chimeric protein construction were applied to trace back the evolution of putrescine-N-methyltransferase activity from spermidine synthase.

The Ding *et al.* [60] investigation was based upon the phylogenetic relationships of species belonging to the family Cruciferae by comparison of the polyamine synthetic genes. A phylogenetic tree was constructed based on the alignment of nucleotide sequences from Nei's genetic distances (a mathematical formula which assumes that genetic differences arise due to mutations and genetic drift). This research elucidated the phylogenetic relationship among crucifer species at the nucleotide level based on the polyamine synthetic gene sequences.

3.11 Polyamines, Health & Therapies

It is now some 30 years since the first observations linked overproduction of polyamines to cancer [61]. The subject of cancer and polyamines has been extensively reviewed [41, 62]. Increased polyamine synthesis and inflammation have long been associated with intraepithelial neoplasia, which are risk factors for cancer development in humans [63]. It is believed that polyamine levels are increased in cancer patients due to the increased growth rate observed in a tumour.

Simple, non-invasive techniques can test for polyamine levels in patients. There have been several examples of these concentrations relating to tumour progression and remission [64]. The polyamine levels in patient serum is a useful tool in assessment of the status of breast cancer patients [65]. Patients in cancer remission will have a urinary polyamine output within the normal range, but that if recurrence of disease takes place, then the output of polyamines will rise [6].

Selective inhibitors have been developed for literally all enzymes of polyamine metabolism, but these polyamine pathway inhibitors are not efficient tumour growth inhibitors. This is due to several reasons; the concentration of drug required can be toxic to humans and that the cells are able to increase polyamine uptake in response to the drug [64]. Low plating efficiency of cell lines and false positive results have made polyamine pathway inhibitors hard to study [64].

Polyamine therapy has also been reported as a possible treatment for other diseases, such as *Pneumocystis* pneumonia, the leading opportunistic disease in AIDS patients. The polyamine pathway of the causative agent can be inhibited, whilst the human polyamine biosynthetic pathway remains unaffected. Partial success has been seen using polyamine inhibitors, this suggests a combination therapy is perhaps a viable option [66].

As already discussed, polyamines are omnipresent in nearly all cells, where they play a fundamental role in cell proliferation and many signalling pathways through their effects on protein kinases, nucleotide cyclases and receptors and through their regulation of the expression of proteins involved in these processes [67]. Therefore, designing polyamine analogues, which can be taken up by the cell using natural transport systems, and then render the internal proteins inactive, is an area for further research.

There are several reported examples of the polyamine pathway being disrupted by specific drugs in a pathogenic micro-organisms with lethal consequences. These therapies work because often the host has a more complex, or different, polyamine pathway which remains unaffected by the drug. The most notable drug of this kind is D,L-α-difluoromethylornithine, commonly known as DFMO.

3.12 DFMO as a Polyamine Inhibitor

D,L-α-difluoromethylornithine (DFMO) is an enzyme-activated, irreversible, inhibitor of the ornithine decarboxylase enzyme (ODC) (Figure 3.3). The specific substrate-based suicide inhibitor binds irreversibly to the alanine racemase (AR) fold of ODC. The AR fold ODC was until recently regarded as

the eukaryotic form of ODC [68]. Some species have the alternative aspartate aminotransferase (AAT) fold ODC, and DFMO is not active against it. Lysine 69 and cysteine 360 in the ODC form covalent adducts with the DFMO. DFMO is cleaved by the ODC, but the product is not released, remaining in the active site and rendering the enzyme irreversibly inactive [69]. Both the L- and D-enantiomers of DFMO can irreversibly inactivate ODC, however the L-form has a 20-times greater probability of forming the complex than the D-form [70]. The inhibition of ODC interferes with the growth of many organisms; including parasitic protozoa [18], bacteria [71], tobacco [19], fungus [72] and mammalian cells [73].

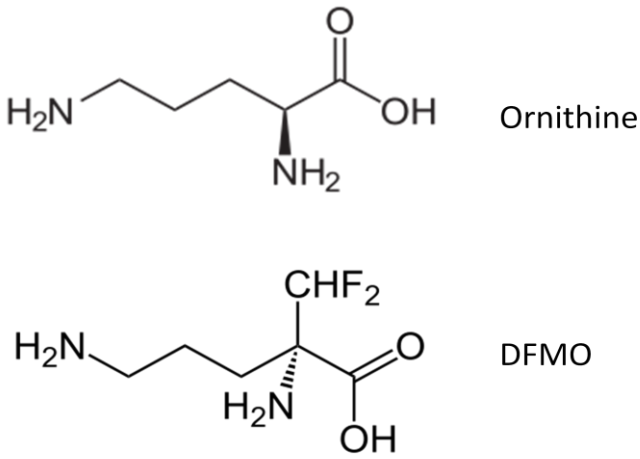


Figure 3.3 The chemical structures of DFMO and ornithine.
DFMO = ($C_6H_{12}F_2N_2O_2$ (S)-2,5-diamino-2-(difluoromethyl)pentanoic acid). Ornithine is the ODC substrate, DFMO is the ODC inhibitor.

DFMO is an effective drug which is currently in clinical use for the treatment of the deadly disease, African Sleeping Sickness, caused by a trypanosome, *Trypanosoma brucei*. The drug has lethal consequences for the trypanosome, but the human patient is unaffected because they have an alternative biosynthetic polyamine pathway (arginine decarboxylase) which is not inhibited by DFMO. The patient recovers completely [17-20].

The drug was initially developed as a cancer medication [61, 62, 64]. Upon topical application, the drug has been shown to be an effective hair growth retardant and is sold under the brand name Vaniqa.

3.13 Bacterial Models for Polyamine Investigations

3.13.1 *Campylobacter jejuni*

Campylobacter, a member of the ε-proteobacteria, belongs to a genus of bacteria that are Gram-negative, spiral and microaerophilic. The strain used in this investigation was 81116 (NCTC 11828); a motile strain [74]. *Campylobacter jejuni* is recognised as one of the main causes of bacterial food borne disease in many developed countries.

There has been very little investigation into the polyamine metabolism in *C. jejuni*. Hamana *et al.* [75] reported that Campylobacter scarcely contained any polyamines, and Forouhar *et al.* [76] have crystallised the biosynthetic arginine decarboxylase (ADC) enzyme. The polyamine pathway of *C. jejuni* has been predicted by comparison of the genome to known polyamine biosynthetic genes (Figure 3.4). The bioinformatic analysis for the role prediction of these genes was performed by Tony Michael and some of the homology of these open reading frames to known genes of the same function was low. In particular, C8J_0166 (CANSDH, Cj0172c) had the lowest homology. Therefore, biochemical testing is required, in conjunction with bioinformatic analysis to unequivocally determine the roles of these genes.

C. jejuni was chosen as a model organism for the ε-proteobacteria because the polyamine profile of this species had not been investigated and *C. jejuni* can be genetically manipulated in the laboratory. The pathway is thought to contain several interesting polyamine-synthetic enzymes. Although polyamine investigations on *C. jejuni* have not been undertaken, another organism from the ε-proteobacteria, *Helicobacter pylori,* has been studied.

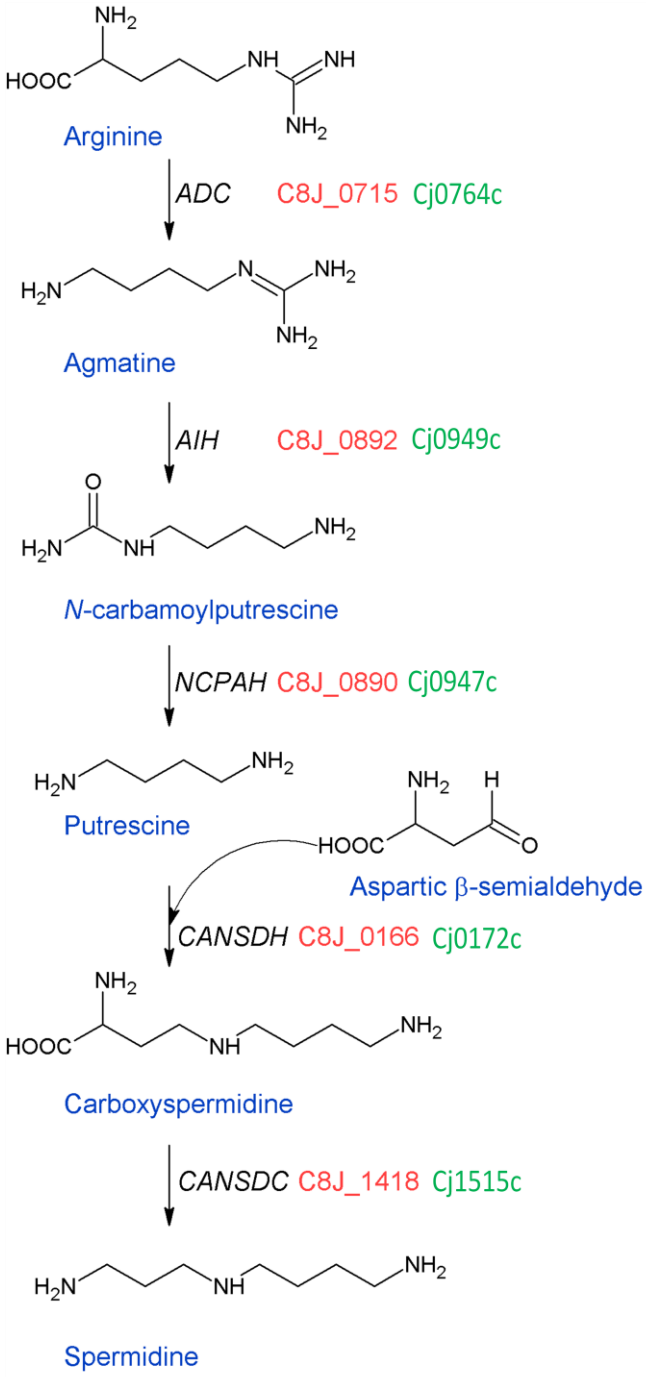


Figure 3.4 The predicted putative spermidine biosynthesis pathway of *C. jejuni* 81116.
Blue, chemical names; *red*, candidate *C. jejuni* open reading frames encoding enzymes. *ADC*, arginine decarboxylase; *AIH*, agmatine iminohydrolase; *NCPAH*, *N*-carbamoylputrescine amidohydrolase; *CANSDH*, carboxyspermidine dehydrogenase; *CANSDC*, carboxyspermidine decarboxylase; *green*, the gene names from *C. jejuni* NCTC 11168.

H. pylori, the Gram-negative, microaerophilic bacterium with strong links to the development of duodenal, gastric ulcers and stomach cancer, has been used in experimentation regarding gastric polyamine levels [77]. The HP0832 (*speE*) gene of *H. pylori* strain 26695 encodes a spermidine synthase, which belongs to the polyamine biosynthetic pathway. Spermidine synthase catalyses the production of spermidine from putrescine and decarboxylated *S*-adenosylmethionine (dcSAM), which serves as an aminopropyl donor. *H. pylori* strain 26695 contains putrescine and spermidine at a molar ratio of 1:3, but no detectable spermine or norspermidine was observed, suggesting that the spermidine biosynthetic pathway may provide the main polyamines in this strain [78]. Methylglyoxal bis(cyclopentylamidino-hydrazone) (MGBCP), a multi-enzyme inhibitor of polyamine biosynthesis inhibits the cell growth of *H. pylori* in a dose-dependent manner. The inhibition was partially reversed by the addition of spermidine [79]. These data suggest the essential nature of polyamines in *H. pylori*.

Both *C. jejuni* (by prediction) and *H. pylori* (by experimentation) contain spermidine as their triamine, which is most likely their sole-triamine and the physiologically relevant polyamine. *C. jejuni* does not contain a spermidine synthase homologue, but instead has carboxyspermidine decarboxylase homologue (C8J_1418). The variety of polyamine biosynthesis between two closely related organisms [80] is of particular interest as we strive to understand the complexity and evolution of the polyamine pathway. This is why it is necessary to investigate this particular organism.

3.13.2 *Rhizobium leguminosarum*

Rhizobium leguminosarum, the α-Proteobacterium, is a nitrogen-fixing plant symbiont. In this study, the sequenced biovar *viciae* strain 3841 was used, a spontaneous streptomycin-resistant mutant of strain 300 [81]. Bacteria of the genus *Rhizobium* play a very important role in agriculture by inducing nitrogen-fixing nodules on the roots of legumes such as peas, beans, clover and alfalfa. This symbiosis can relieve the requirements for added nitrogenous fertiliser during the growth of leguminous crops.

Ornithine decarboxylase (ODC) catalyses the first and rate-limiting step in the biosynthesis of polyamines in most eukaryotes [58]. It catalyses the decarboxylation of ornithine to produce the diamine putrescine and is one of the most highly regulated enzymes of eukaryotic organisms [82]. ODC has been detected in a range of prokaryotic organisms [83].

The polyamine production pathway in *R. leguminosarum* is relatively simple; ODC produces putrescine from ornithine, then homospermidine synthase generates homospermidine from two molecules of putrescine (Figure 3.5). Homospermidine, the relatively uncommon polyamine, is the only triamine in *R. leguminosarum* [84]. The relatively simple pathway of *R. leguminosarum* makes it an ideal model organism for polyamine investigation and experimentation in the α-Proteobacteria.

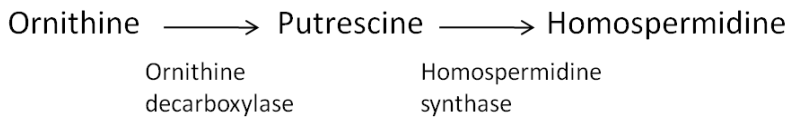


Figure 3.5 The predicted polyamine pathway of *R. leguminosarum.*
One molecule of ornithine forms one putrescine, then two putrescine combine to form one homospermidine (as depicted in Figure 3.6).

3.14 Homospermidine & Homospermidine Synthase

Homospermidine is a linear triamine which contains eight carbon and three amine groups (Figure 3.2). Initially, the compound was discovered as the linked backbone component of five isolated alkaloids of the Solanaceae [85]. Following this discovery, homospermidine was identified as a free polyamine in the leaves of the sandalwood tree [86], and was later recognised in root nodules and in the Rhizobia spp. [87].

3.14.1 Homospermidine Distribution

Homospermidine has been detected in widely diverse organisms [41]. Although homospermidine is generally considered to be an uncommon polyamine, it is found in nitrogen-fixing Cyanobacteria [88], in the Chloroflexi

and Chlorobi [89], in the Flavobacteriaceae, Flexibacteriaceae, and Sphingobacteriaceae of the Bacteroidetes [90, 91], in the Thermus phylum [92] and in the Gemmatimonadetes, Acidobacteria and Planctomycetes [92].

Within the Proteobacteria, homospermidine has been found in 18 out of 35 analysed genera of α-Proteobacteria; it is widespread in the δ-Proteobacteria, absent in the ε-Proteobacteria, and prominent in the pathogenic γ-proteobacterial genus Legionella [93]. As homospermidine has never been associated with the ε-Proteobacteria, *C. jejuni* was studied as a model organism to see if the homospermidine can be utilised by the bacteria (discussed in Chapter 5).

Homospermidine appears to be absent from the Aquificae, Thermotogae, Firmicutes and Actinobacteria [92]. Within the archaea, homospermidine is found in methanogenic archaea [94]. Among single-cell eukaryotes, homospermidine has also been found in the slime mould Physarum [95]. Homospermidine has also been identified in the plant kingdom. Certain plant lineages, 29 different species randomly selected from 18 angiosperm families as well as a few other terrestrial plant species, all produce homospermidine for pyrrolizidine alkaloid biosynthesis [96].

3.14.2 Homospermidine Synthase

One enzyme responsible for homospermidine production is homospermidine synthase (HSS). A HSS enzymatic activity was first reported by Tait [97] in the α-Protobacterium *Rhodopseudomonas viridis* (later renamed *Blastochloris viridis*). Homospermidine can also be produced by another enzyme – deoxyhypusine synthase-like-homospermidine synthase (DHS-HSS) – which is discussed later in this chapter (Section 3.14.7) [96]. Homospermidine synthase enzymes are found in both plants and bacteria [98].

3.14.3 Homospermidine Synthase Mechanism

A diagram showing the simplified mechanism of HSS is shown in Figure 3.6.

Tait [97] showed HSS activity when cell-free extracts of *R. viridis* formed homospermidine when incubated with putrescine and NAD^+. The reaction was described as follows: the conversion was effected by a single enzyme that oxidatively deaminated one molecule of putrescine in the presence of NAD^+ to yield ammonia, enzyme-bound γ-aminobutyraldehyde, and enzyme-bound NADH. A Schiff-base complex was then formed with a second molecule of putrescine and the enzyme-bound γ-aminobutyraldehyde, and this complex was then reduced by the enzyme-bound NADH to form homospermidine, regenerating NAD^+.

When Yamamoto *et al.* [99] investigated HSS in *Acinetobacter tartarogenes* they concluded the mechanism to be similar to that described by Tait [97]. The enzyme functions by two serial reductions; firstly, putrescine is oxidised to form 4-aminobutyraldehyde, and then, a subsequent reduction of the putative Schiff base formed between this aldehyde and a second molecule of putrescine to homospermidine. Yamamoto also concluded HSS will not function upon diaminopropane and cadaverine when used as substrates. In addition, $NADP^+$ could not be substituted for the true co-factor NAD^+ [99].

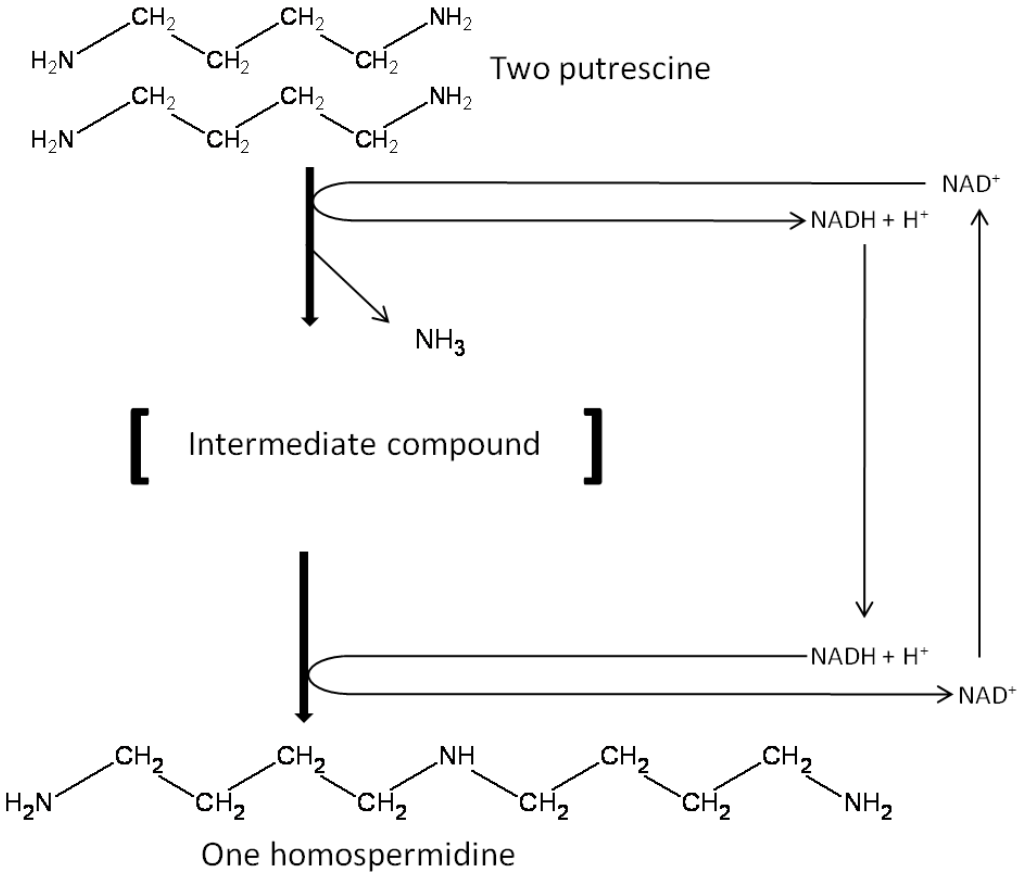


Figure 3.6 The formation of homospermidine (adapted from Tholl [100])

Tholl *et al.* [100] investigated the reaction mechanism of the HSS extracted from *R. viridis.* It was suggested that the enzymatic activity was based on the existence of two binding sites at the catalytic centre for the two substrate molecules. One of these sites is thought to be less substrate specific (i.e. will bind homologous diamines such as diaminopropane, cadaverine, diaminohexane and spermidine) and the other is specific to putrescine. Tholl's evaluation that diaminopropane could be used as a substrate of HSS disputed Yamamoto's earlier work.

The idea of HSS substrate specificity is relevant to the experiments performed during this investigation. The possibility of a two-site enzyme is discussed when HSS is expressed in the *E. coli* host, because the natural polyamines found in the *E. coli* may be used as a substrates for the non-native HSS enzyme.

Böttcher *et al.* [98] studied the differences and similarities between plant and bacterial HSSs. It was seen that plant HSS (taken from *Eupatorium*

cannabinum) would only ever produce homospermidine even when a variety of substrates were provided to the enzyme *in vitro*. The bacterial HSS (from *R. viridis* again) was able to produce alternative minor products when supplied with artificial substrates in the enzyme assay (Table 3.1).

Substrates	Products Observed
DAP	-
PUT	HSPD
CAD	-
PUT & DAP	HSPD & SPD
PUT & CAD	HSPD & ABC
PUT & DAHex	HSPD & 4-aminobutyl-1,6-DAHex
PUT & DAHep	HSPD & 4-aminobutyl-1,7-DAHep(trace)
SPD	HSPD, DAP & PUT
SPD & PUT	HSPD & DAP
SPD & CAD	HSPD, ABC & DAP
SPD & DAHex	HSPD, 4-aminobutyl-1,6-DAHex & DAP

Table 3.1 The products produced by bacterial (*R. viridis*) homospermidine synthase *in vitro* with the addition of various substrates.
DAP = Diaminopropane, PUT = Putrescine, CAD = Cadaverine, ABC = Aminobutylcadaverine, HSPD = Homospermidine, SPD = Spermidine, DAHex = Diaminohexane, DAHep = Diaminoheptane [98].

3.14.4 Homospermidine Synthase Cloning, Expression & Characterisation

Decades after the first formation of homospermidine was described in *R. viridis* [97], the gene encoding the HSS enzyme of *R. viridis* was cloned and the encoded enzyme was a homodimer with subunits of 52 kDa [100]. A viral orthologue of this gene from the large *Paramecium bursaria* Chlorella virus 1 was shown to encode a functional HSS [13]. The kinetic parameters of HSS have been reported from purified extracts of the γ-proteobacterium *A. tartarogenes* [99] and α-Protobacterium *R. viridis* [98] (Table 3.2). HSS from *A. tartarogenes* and *R. viridis* were reported to be similar. They both appear to be of the same molecular weight, with similar sized subunits. The optimum temperature, pH and co-factor concentrations are similar for both enzymes.

These data may suggest that the enzymes have been closely conserved during evolution. This argues a specific function for the product, homospermidine. However, the investigations to date have not presented an implication or function for homospermidine.

Property	***A. tartarogenes***	***R. viridis***
Native M_r	102	100
Subunit M_r	52	52
Optimum K^+ (mM)	50	40
Optimum DTT (mM)	1	2
Optimum pH	8.7	8.8
Optimum temperature (^oC)	30	37

Table 3.2 The reported properties of the previously studied HSS enzymes from *Acinetobacter tartarogenes* and *Rhodopseudomonas viridis* (adapted from Yamamoto *et al.* [99] and supplemented with data from Tholl *et al.* [100]).

3.14.5 Homospermidine Synthase Phylogenetic Distribution

In the field of evolutionary biological classification, genetic relationships can be predicted using the structure and distribution of certain compounds, such as polyamines, among the organisms being classified. Homospermidine has been used as a chemotaxonomic marker in certain bacterial taxa, such as the α-Proteobacteria [101].

Homospermidine synthase, has been identified in several species and shows a high percentage of conservation between organisms. BLAST searches have enabled genes in sequenced organisms to be identified as possible *hss*. The organisms identified are located in phylogenetically diverse locations and to illustrate this point they have been highlighted in Figure 3.7. Organisms with possible *hss* include *Methanosarcina mazei, Pelodictyon phaeoclathratiforme, Magnetospirillum magneticum* and *Paramecium tetraurelia* to name a few. The number of organisms with possible *hss* genes is increasing rapidly as the number of sequenced genomes increases.

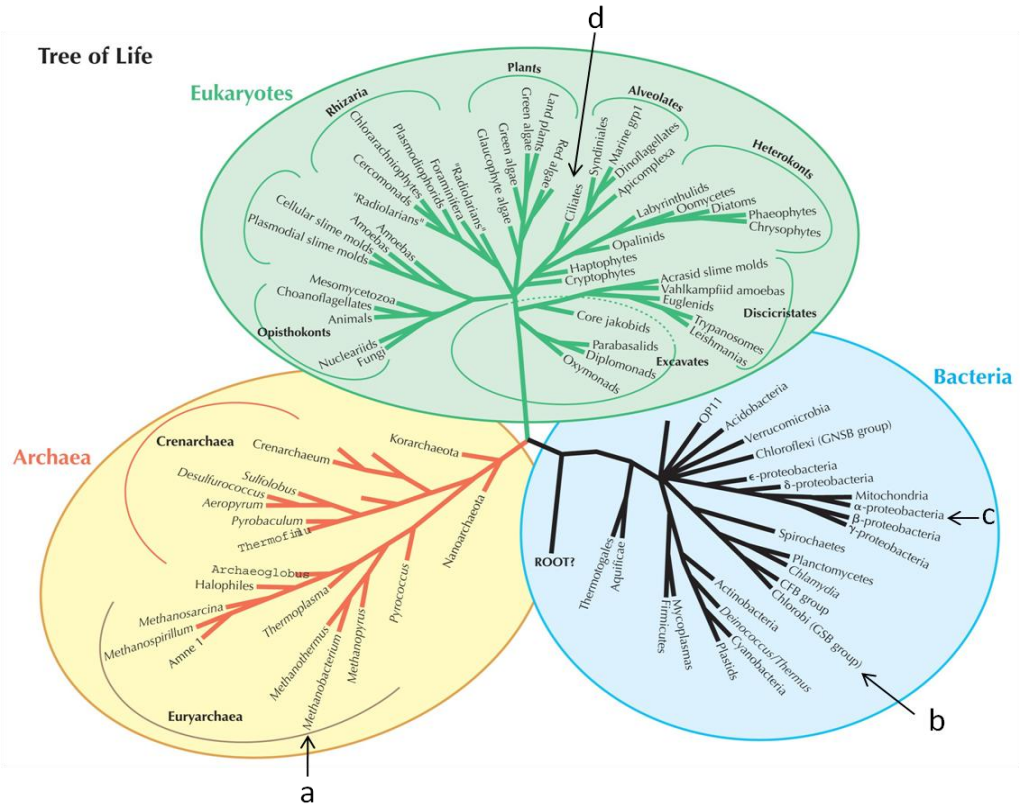


Figure 3.7 'The Tree of Life' adapted from Eisen [102] to highlight four diverse species containing *hss*.
Methanosarcina mazei (a), *Pelodictyon phaeoclathratiforme* (b), *Magnetospirillum magneticum* (c) and *Paramecium tetraurelia* (d).

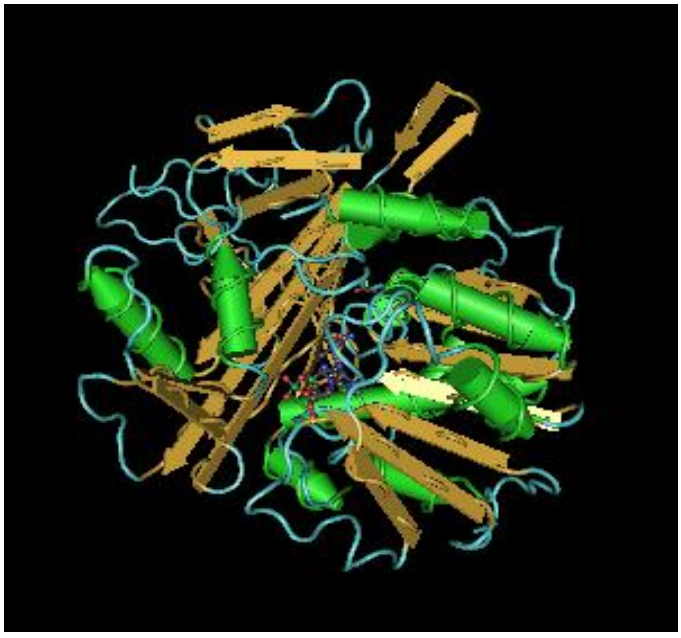


Figure 3.8 Crystallised homospermidine synthase from *Legionella pneumophila* [103].

3.14.6 Homospermidine Synthase Structure

Homospermidine synthase from *Legionella pneumophila subsp. pneumophila* str. Philadelphia 1 has been crystallised and the structure has not been published but is available from GenBank (Figure 3.8) [103]. The crystal structure was formed using x-ray diffraction, with a resolution of 2.5 Å, and was shown to be in complex with NAD^+. This agrees with the predicted mechanism of HSS (Figure 3.6); where NAD^+ is used as a recycled co-factor and binds to the enzyme. Unfortunately, no further structural investigation has been performed on the data generated by this investigation.

3.14.7 An Alternative Homospermidine Synthase

Homospermidine is synthesised in some plants as the precursor for pyrrolizidine alkaloid production [96]. In contrast to the HSS mechanism already discussed, the plant HSS is derived from deoxyhypusine synthase (DHS), an enzyme that transfers an aminobutyl group from spermidine to a unique lysine residue in translation initiation factor eIF5A [104, 105] and is unrelated to the bacterial HSS enzyme described above. The DHS-like HSS has an absolute requirement for spermidine as a substrate, and the aminobutyl group from spermidine is transferred to a molecule of putrescine to form homospermidine [98].

Genetic studies indicate that some bacteria which produce homospermidine, use an enzyme derived from DHS. This enzyme has evolved and lost its major function (hypusine formation), but still functions readily as a HSS. Oshima [106] predicts that the DHS-like HSS gene found in *Thermus thermophilius*, an extreme thermophile, can form homospermidine from agmatine. Although, they failed to show the activity under *in vitro* conditions.

3.14.8 Homospermidine Function

A specific intracellular or molecular function of homospermidine in bacteria has yet to be identified. However, homospermidine is an essential precursor in the biosynthesis of plant defence compounds, pyrrolizidine alkaloids [107].

Many eubacteria, particularly those which can fix nitrogen or form nitrogen fixing nodules, have been tested for the presence of homospermidine. Although several species of nitrogen fixing bacteria have been found to contain homospermidine, such as cyanobacteria*, Rhodospirillum* and *Chromatium*, other nitrogen fixing bacteria, such as *Azotobacter,* do not contain homospermidine [13, 94, 101, 108, 109]. There is no evidence for a biochemical relationship between the polyamine homospermidine and the ability to fix atmospheric nitrogen.

There are organisms which contain only homospermidine or spermidine as their triamine. The very existence of organisms that use only the asymmetrical spermidine or only the symmetrical homospermidine as their sole triamine has profound implications for the molecular mechanisms of action of polyamines. Spermidine binds to RNA and influences ribosome function and mRNA translation [110] and also binds to DNA, thereby influencing DNA structure and compaction [111]. It has been suggested that the manner of binding of polyamines to nucleic acids is specific; however, if cells cope with only spermidine or only homospermidine, the exact nature of the polyamine structure and mode of binding would not appear to be critical for function.

New polyamines and new homospermidine derivatives, including long linear and branched pentamine, hexamines and heptamines, are constantly being discovered. In addition to homopentamine and homohexamine, some other eight tertiary branched homospermidine derived polymers have been identified [41]. This indicates that the homospermidine molecule is an important intermediate in biosynthetic and metabolic pathways, but the variety of these homospermidine derivatives suggests that, in these organisms, homospermidine does not have a particular role and is merely an intermediate. However, in bacterial species where homospermidine is the main triamine, it is almost certain that it has a function, which we are yet to elucidate.

3.15 Conclusion

Almost all organisms require polyamines to live, but we know there is no single 'universal' polyamine. A specific polyamine may be essential in one species, but toxic to another. A polyamine may have a defined cellular function in one micro-organism, but a different function in another. The field of polyamine study is vast, and there are still many aspects of these amines not fully understood.

Polyamine metabolism is a complicated and delicate process, the differences and similarities between various species and phyla have been examined and documented. The complexity and multiplicity of the polyamine pathway has been studied intensely, as this pathway is used for examining the evolution of metabolic diversity in organisms. However, the exact role of some cellular polyamines is still unknown and new discoveries within the polyamine pathway are constantly being revealed. Polyamine concentration can also be used to detect illness, infection or chemical imbalance.

Polyamines are essential for life and serve a variety of cellular functions, including osmoregulation, phospholipid biosynthesis, cell migration, cell proliferation and maintenance of the macromolecular structure. Depletion of polyamines dramatically alters the cellular functions on many levels, both beneficially and detrimentally. Cellular aspects which are altered include the arrangement of the cytoskeleton, receptor functions, the activities of signal transduction proteins, the levels of several protooncogenes and the expression and concentration of growth factors.

Once an understanding of how polyamines function and interact within the cell is fully elucidated, this information can be related to a variety of medical and ecological applications. If cells can be manipulated by alteration of polyamine content, or modification of enzymes required in the polyamine pathway, then pathogenic organisms can be eradicated and useful organisms can be cultured.

This thesis will investigate the evolution and phylogenetic distribution of the HSS gene. In addition, the HSS enzyme will be kinetically characterised and

compared to those which have previously been characterised. The substrate specificity of HSS will be explored and the ability of micro-organisms to use alternative triamines will be investigated.

Chapter Three - References

1. Tabor CW, Tabor H: **Polyamines in microorganisms**. *Microbiological Reviews* 1985, **49**(1):81-99.
2. Pegg AE: **Recent Advances in the Biochemistry of Polyamines in Eukaryotes**. *Biochemical Journal* 1986, **234**(2):249-262.
3. Evans PT, Malmberg RL: **Do Polyamines Have Roles in Plant Development?** *Annual Review of Plant Physiology and Plant Molecular Biology* 1989, **40**(1):235-269.
4. Edreva A: **Polyamines in Plants**. *Journal of Plant Physiology* 1996, **22**(1-2):73-101.
5. Montanez R, Sanchez-Jimenez F, Aldana-Montes JF, Medina MA: **Polyamines: metabolism to systems biology and beyond**. *Amino Acids* 2007, **33**(2):283-289.
6. Wallace HM, Fraser AV, Hughes A: **A perspective of polyamine metabolism**. *Biochemistry Journal* 2003, **376**(1):1-14.
7. Tobias KE, Kahana C: **Intersubunit location of the active site of mammalian ornithine decarboxylase as determined by hybridization of site-directed mutants**. *Biochemistry* 1993, **32**(22):5842-5847.
8. Hayashi S, Murakami Y: **Rapid and regulated degradation of ornithine decarboxylase**. *Biochemistry Journal* 1995, **306 (1)**:1-10.
9. Hamana K, Matsuzaki S: **Polyamines as a chemotaxonomic marker in bacterial systematics**. *Crit Rev Microbiol* 1992, **18**(4):261-283.
10. Coffino P: **Regulation of cellular polyamines by antizyme**. *Nature Reviews Molecular Cell Biolology* 2001, **2**(3):188-194.
11. Oshima T: **Novel polyamines of extremely thermophilic bacteria**. In: *The Physiology of Polyamines.* Edited by Bachrach U, Heimer YM. New York; 1978: 211-220.
12. Oshima T: **A new polyamine, thermospermine, 1,12-diamino-4,8-diazadodecane, from an extreme thermophile**. *Journal of Biological Chemistry* 1979, **254**(18):8720-8722.
13. Kaiser A, Vollmert M, Tholl D, Graves MV, Gurnon JR, Xing W, Lisec AD, Nickerson KW, Van Etten JL: **Chlorella virus PBCV-1 encodes a functional homospermidine synthase**. *Virology* 1999, **263**(1):254-262.

14. Hamana K, Matsuzaki S, Sakakibara M: **Distribution of Sym-Homospermidine in Eubacteria, Cyanobacteria, Algae and Ferns**. *Fems Microbiology Letters* 1988, **50**(1):11-16.
15. Kuehn GD, Rodriguez-Garay B, Bagga S, Phillips GC: **Novel Occurrence of Uncommon Polyamines in Higher Plants**. *Plant Physiology* 1990, **94**(3):855-857.
16. Hamana K, Niitsu M, Matsuzaki S, Samejima K, Igarashi Y, Kodama T: **Novel linear and branched polyamines in the extremely thermophilic eubacteria Thermoleophilum, Bacillus and Hydrogenobacter**. *Biochemical Journal* 1992, **284 (3)**:741-747.
17. Wallace HM, Fraser AV: **Inhibitors of polyamine metabolism: review article**. *Amino Acids* 2004, **26**(4):353-365.
18. Carrillo C, Cejas S, Cortes M, Ceriani C, Huber A, Gonzalez NS, Algranati ID: **Sensitivity of *trypanosomatid protozoa* to DFMO and metabolic turnover of ornithine decarboxylase**. *Biochemical and Biophysical Research Communications* 2000, **279**(2):663-668.
19. Slocum RD, Bitonti AJ, McCann PP, Feirer RP: **DL-alpha-difluoromethyl[3,4-3H]arginine metabolism in tobacco and mammalian cells. Inhibition of ornithine decarboxylase activity after arginase-mediated hydrolysis of DL-alpha-difluoromethylarginine to DL-alpha-difluoromethylornithine**. *Biochemistry Journal* 1988, **255**(1):197-202.
20. Schwartz B, Hittelman A, Daneshvar L, Basu HS, Marton LJ, Feuerstein BG: **A new model for disruption of the ornithine decarboxylase gene, SPE1, in *Saccharomyces cerevisiae* exhibits growth arrest and genetic instability at the MAT locus**. *Biochemistry Journal* 1995, **312** (1):83-90.
21. Chattopadhyay MK, Murakami Y, Matsufuji S: **Antizyme regulates the degradation of ornithine decarboxylase in fission yeast *Schizosaccharomyces pombe*. Study in the spe2 knockout strains**. *Journal of Biological Chemistry* 2001, **276**(24):21235-21241.
22. Jin Y, Bok JW, Guzman-de-Pena D, Keller NP: **Requirement of spermidine for developmental transitions in *Aspergillus nidulans***. *Molecular Microbiology* 2002, **46**(3):801-812.
23. Guevara-Olvera L, Xoconostle-Cazares B, Ruiz-Herrera J: **Cloning and disruption of the ornithine decarboxylase gene of *Ustilago***

***maydis*: evidence for a role of polyamines in its dimorphic transition**. *Microbiology* 1997, **143 (Pt 7)**:2237-2245.

24. Li F, Hua SB, Wang CC, Gottesdiener KM: **Procyclic *Trypanosoma brucei* cell lines deficient in ornithine decarboxylase activity**. *Molecular and Biochemical Parasitology* 1996, **78**(1-2):227-236.
25. Jiang Y, Roberts SC, Jardim A, Carter NS, Shih S, Ariyanayagam M, Fairlamb AH, Ullman B: **Ornithine decarboxylase gene deletion mutants of *Leishmania donovani***. *Journal of Biological Chemistry* 1999, **274**(6):3781-3788.
26. Gillin FD, Reiner DS, McCann PP: **Inhibition of growth of *Giardia lamblia* by difluoromethylornithine, a specific inhibitor of polyamine biosynthesis**. *Journal of Protozoology* 1984, **31**(1):161-163.
27. Pendeville H, Carpino N, Marine JC, Takahashi Y, Muller M, Martial JA, Cleveland JL: **The ornithine decarboxylase gene is essential for cell survival during early murine development**. *Molecular and Cellular Biology* 2001, **21**(19):6549-6558.
28. Nishimura K, Nakatsu F, Kashiwagi K, Ohno H, Saito T, Igarashi K: **Essential role of S-adenosylmethionine decarboxylase in mouse embryonic development**. *Genes Cells* 2002, **7**(1):41-47.
29. Urano K, Hobo T, Shinozaki K: **Arabidopsis ADC genes involved in polyamine biosynthesis are essential for seed development**. *FEBS Letters* 2005, **579**(6):1557-1564.
30. Hafner EW, Tabor CW, Tabor H: **Mutants of *Escherichia coli* that do not contain 1,4-diaminobutane (putrescine) or spermidine**. *Journal of Biological Chemistry* 1979, **254**(24):12419-12426.
31. Tabor H, Hafner EW, Tabor CW: **Construction of an *Escherichia coli* strain unable to synthesize putrescine, spermidine, or cadaverine: characterization of two genes controlling lysine decarboxylase**. *Journal of Bacteriology* 1980, **144**(3):952-956.
32. Flores H, Protacio C, Signs M: **Primary and secondary metabolism of polyamines in plants**. In: *Plant Nitrogen Metabolism.* Edited by Poulton J, Romeo J, Conn E, vol. 23. New York: Plenum Press; 1989
33. Roy M, Ghosh B: **Polyamines, both common and uncommon, under heat stress in rice (*Oryza sativa*) callus.** *Physiologia Plantarum* 1996, **98**(1):196-200.

34. Huang Y, Marton LJ, Woster PM, Casero RA: **Polyamine analogues targeting epigenetic gene regulation**. *Essays in Biochemistry* 2009, **46**:95-110.

35. Iacomino G, Picariello G, Sbrana F, Di Luccia A, Raiteri R, D'Agostino L: **DNA is Wrapped by the Nuclear Aggregates of Polyamines: The Imaging Evidence**. *Biomacromolecules* 2011.

36. Schuber F: **Influence of polyamines on membrane functions**. *Biochemistry Journal* 1989, **260**(1):1-10.

37. delaVega AL, Delcour AH: **Cadaverine induces closing of E. coli porins**. *EMBO J* 1995, **14**(23):6058-6065.

38. Marton L, Morris D: **Molecular and cellular function of the polyamines**. In: *Inhibition of Polyamine Metabolism.* Edited by McCann P, Pegg A, Sjoerdsma A. San Diego: Academic Press.; 1987: 79-105.

39. Tang W, Newton RJ: **Polyamines promote root elongation and growth by increasing root cell division in regenerated Virginia pine (*Pinus virginiana Mill.*) plantlets**. *Plant Cell Reports* 2005, **24**(10):581-589.

40. Tadolini B, Cabrini L, Landi L, Varani E, Pasquali P: **Polyamine binding to phospholipid vesicles and inhibition of lipid peroxidation**. *Biochemical and Biophysical Research Communications* 1984, **122**(2):550-555.

41. Cohen S: **A guide to the Polyamines**. New York: Oxford University Press; 1998.

42. Munro GF, Bell CA: **Phospholipid turnover in a conditional polyamine auxotroph of *Escherichia coli***. *Journal of Bacteriology* 1973, **116**(3):1479-1481.

43. Gerlach GF, Clegg S, Ness NJ, Swenson DL, Allen BL, Nichols WA: **Expression of type 1 fimbriae and mannose-sensitive hemagglutinin by recombinant plasmids**. *Infect Immun* 1989, **57**(3):764-770.

44. Tabor CW, Tabor H: **1,4-Diaminobutane (putrescine), spermidine, and spermine**. *Annual Review of Biochemistry* 1976, **45**:285-306.

45. Kusano T, Berberich T, Tateda C, Takahashi Y: **Polyamines: essential factors for growth and survival**. *Planta* 2008, **228**(3):367-381.

46. Kawahara H, Inagaki K, Kishimoto N, Sugio T, Tano T: **Effects of polyamine on the growth of *Acidiphilium facilis* 24R under acidic conditions**. *Agricultural and Biological Chemistry* 1991, **55**:679-685

47. Friedman ME, Bachrach U: **Inhibition of protein synthesis by spermine in growing cells of *Staphylococcus aureus***. *Journal of Bacteriology* 1966, **92**(1):49-55.

48. Mager J, Traub A, Grossowk ZN: **Cultivation of Pasteurella tularensis in chemically defined media: effect of buffers and spermine**. *Nature* 1954, **174**(4433):747-748.

49. Yoshida M, Kashiwagi K, Shigemasa A, Taniguchi S, Yamamoto K, Makinoshima H, Ishihama A, Igarashi K: **A unifying model for the role of polyamines in bacterial cell growth, the polyamine modulon**. *Journal of Biological Chemistry* 2004, **279**(44):46008-46013.

50. Raina A, Jansen M, Cohen SS: **Polyamines and the accumulation of ribonucleic acid in some polyauxotrophic strains of *Escherichia coli***. *Journal of Bacteriology* 1967, **94**(5):1684-1696.

51. Schipper RG, Penning LC, Verhofstad AA: **Involvement of polyamines in apoptosis. Facts and controversies: effectors or protectors?** *Seminars in Cancer Biology* 2000, **10**(1):55-68.

52. Stevens L, Pascoe G: **The location of spermine in bacterial ribosomes as indicated by 1,5-difluoro-2,4-dinitrobenzene and by ethidium bromide**. *Biochemistry Journal* 1972, **128**(2):279-289.

53. Rubin RL: **Spermidine-Deoxyribonucleic acid interaction *in vitro* and in *Escherichia coli***. *Journal of Bacteriology* 1977, **129**(2):916-925.

54. Basu HS, Wright WD, Deen DF, Roti-Roti J, Marton LJ: **Treatment with a polyamine analog alters DNA-matrix association in HeLa cell nuclei: a nucleoid halo assay**. *Biochemistry* 1993, **32**(15):4073-4076.

55. Kleppe K, Osland A, Fosse V, Male R, Lossius I, Helland D, Lillehaug JR, Raae AJ, Kleppe RK, Nes IF: **Effect of polyamines on enzymes involved in DNA repair**. *Biology and Medicine Journal* 1981, **59**(5-6):374-380.

56. Hamana K, Matsuzaki S: **Polyamines as a chemotaxonomic marker in bacterial systematics**. *Critical Reviews in Microbiology* 1992, **18**(4):261-283.

57. Tomkins JP, Wood TC, Stacey MG, Loh JT, Judd A, Goicoechea JL, Stacey G, Sadowsky MJ, Wing RA: **A marker-dense physical map of the *Bradyrhizobium japonicum* genome**. *Genome Research* 2001, **11**(8):1434-1440.

58. Ivanov IP, Firth AE, Atkins JF: **Recurrent emergence of catalytically inactive ornithine decarboxylase homologous forms that likely have regulatory function**. *The Journal of Molecular Evolution* 2010, **70**(3):289-302.

59. Biastoff S, Reinhardt N, Reva V, Brandt W, Drager B: **Evolution of putrescine N-methyltransferase from spermidine synthase demanded alterations in substrate binding**. *FEBS Letters* 2009, **583**(20):3367-3374.

60. Ding SL, Lu G, Li JY, Ren Y, Cao JS: **Cloning and evolutionary analysis of homologous sequences of SAMDC gene in Cruciferae**. *Yi Chuan* 2007, **29**(1):109-117.

61. Seiler N: **Thirty Years of Polyamine-Related Approaches to Cancer Therapy. Retrospect and Prospect. Part 1. Selective Enzyme Inhibitors** In: *Current Drug Targets* vol. 4: Bentham Science Publishers; 2003: 537-564.

62. Wallace HM: **The polyamines: past, present and future**. *Essays in Biochemistry* 2009, **46**:1-9.

63. Babbar N, Gerner EW: **Targeting polyamines and inflammation for cancer prevention**. *Recent Results in Cancer Research* 2010, **188**:49-64.

64. Bachrach U: **Polyamines and cancer: minireview article**. *Amino Acids* 2004, **26**(4):307-309.

65. Byun JA, Lee SH, Jung BH, Choi MH, Moon MH, Chung BC: **Analysis of polyamines as carbamoyl derivatives in urine and serum by liquid chromatography-tandem mass spectrometry**. *Biomedical Chromatography* 2008, **22**(1):73-80.

66. Liao CP, Lasbury ME, Wang SH, Zhang C, Durant PJ, Murakami Y, Matsufuji S, Lee CH: **Pneumocystis mediates overexpression of antizyme inhibitor resulting in increased polyamine levels and apoptosis in alveolar macrophages**. *Journal of Biological Chemistry* 2009, **284**(12):8174-8184.

67. Melchiorre C, Bolognesi ML, Minarini A, Rosini M, Tumiatti V: **Polyamines in drug discovery: from the universal template approach to the multitarget-directed ligand design strategy**. *Journal of Medicinal Chemistry* 2010, **53**(16):5906-5914.

68. Shaw FL, Elliott KA, Kinch LN, Fuell C, Phillips MA, Michael AJ: **Evolution and multifarious horizontal transfer of an alternative biosynthetic pathway for the alternative polyamine sym-homospermidine**. *Journal of Biological Chemistry* 2010, **285**(19):14711-14723.

69. Poulin R, Lu L, Ackermann B, Bey P, Pegg AE: **Mechanism of the irreversible inactivation of mouse ornithine decarboxylase by alpha-difluoromethylornithine. Characterization of sequences at the inhibitor and coenzyme binding sites**. *Journal of Biological Chemistry* 1992, **267**(1):150-158.

70. Qu N, Ignatenko NA, Yamauchi P, Stringer DE, Levenson C, Shannon P, Perrin S, Gerner EW: **Inhibition of human ornithine decarboxylase activity by enantiomers of difluoromethylornithine**. *Biochemistry Journal* 2003, **375**(Pt 2):465-470.

71. Barry DP, Asim M, Leiman DA, de Sablet T, Singh K, Casero RA, Chaturvedi R, Wilson KT: **Difluoromethylornithine Is a Novel Inhibitor of Helicobacter pylori Growth, CagA Translocation, and Interleukin-8 Induction**. *PLoS One* 2011, **6**(2):e17510.

72. Pieckenstain FL, Garriz A, Chornomaz EM, Sanchez DH, Ruiz OA: **The effect of polyamine biosynthesis inhibition on growth and differentiation of the phytopathogenic fungus *Sclerotinia sclerotiorum***. *Antonie Van Leeuwenhoek* 2001, **80**(3-4):245-253.

73. Gerner EW, Mamont PS: **Restoration of the polyamine contents in rat hepatoma tissue-culture cells after inhibition of polyamine biosynthesis relationship with cell proliferation**. *European Journal of Biochemistry* 1986, **156**(1):31-35.

74. Pearson BM, Gaskin DJ, Segers RP, Wells JM, Nuijten PJ, van Vliet AH: **The complete genome sequence of *Campylobacter jejuni* strain 81116 (NCTC11828)**. *Journal of Bacteriology* 2007, **189**(22):8402-8403.

75. Hamana K, Saito T, Okada M, Niitsu M: **Polyamine Distribution Profiles among Some Members within Delta-and Epsilon-**

Subclasses of Proteobacteria. *Microbial Culture Collection* 2004, **20**(1):3-8.

76. Forouhar F, Lew S, Seetharaman J, Xiao R, Acton TB, Montelione GT, Tong L: **Structures of bacterial biosynthetic arginine decarboxylases**. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* 2010, **66**(12):1562-1566.
77. Linsalata M, Russo F, Notarnicola M, Berloco P, Di Leo A: **Polyamine profile in human gastric mucosa infected by *Helicobacter pylori***. *Italian Journal of Gastroenterology and Hepatology* 1998, **30**(5):484-489.
78. Lee MJ, Huang CY, Sun YJ, Huang H: **Cloning and characterization of spermidine synthase and its implication in polyamine biosynthesis in *Helicobacter pylori* strain 26695**. *Protein Expression and Purification* 2005, **43**(2):140-148.
79. Takaji S, Hibasami H, Imoto I, Hashimoto Y, Nakao K, Ikemura N, Taguchi Y, Nakashima K: **Growth inhibition of *Helicobacter pylori* by a polyamine synthesis inhibitor, methylglyoxal bis(cyclopentylamidinohydrazone)**. *Letters in Applied Microbiology* 1997, **25**(3):177-180.
80. Eppinger M, Baar C, Raddatz G, Huson DH, Schuster SC: **Comparative analysis of four Campylobacterales**. *Nature Reviews Microbiology* 2004, **2**(11):872-885.
81. Young JP, Crossman LC, Johnston AW, Thomson NR, Ghazoui ZF, Hull KH, Wexler M, Curson AR, Todd JD, Poole PS *et al*: **The genome of *Rhizobium leguminosarum* has recognizable core and accessory components**. *Genome Biology* 2006, **7**(4):R34.
82. Davis RH, Morris DR, Coffino P: **Sequestered end products and enzyme regulation: the case of ornithine decarboxylase**. *Microbiology Reviews* 1992, **56**(2):280-290.
83. Lee J, Michael AJ, Martynowski D, Goldsmith EJ, Phillips MA: **Phylogenetic diversity and the structural basis of substrate specificity in the beta/alpha-barrel fold basic amino acid decarboxylases**. *Journal of Biological Chemistry* 2007, **282**(37):27115-27125.

84. Hamana K, Minamisawa K, Matsuzaki S: **Polyamines in Rhizobium, Bradyrhizobium, Azorhizobium and Argobacterium**. *FEMS Microbiology Letters* 1990, **71**(1-2):71-76.

85. Badawi MM, Bernauer K, van den Broek P, Groger D, Guggisberg A, Johne S, Kompis I, Schneider F, Veith HJ, Hesse M *et al*: **Macrocyclic spermidine and spermine alkaloids**. *Pure Applied Chemistry* 1973, **33**(1):81-107.

86. Kuttan R, Radhakrishnan AN, Spande T, Witkop B: **sym-Homospermidine, a naturally occurring polyamine**. *Biochemistry* 1971, **10**(3):361-365.

87. Smith TA: **Homospermidine in Rhizobium and Legume Root-Nodules**. *Phytochemistry* 1977, **16**(2):278-279.

88. Hamana K, Miyagawa K, Matsuzaki S: **Occurrence of sym-homospermidine as the major polyamine in nitrogen-fixing cyanobacteria**. *Biochemical and Biophysical Research Communications* 1983, **112**(2):606-613.

89. Norgaard E, Sirevag R, Eliassen KA: **Evidence for the occurrence of sym-homospermidine in green phototrophic bacteria** *FEMS Microbiol Letters* 1983, **20**(2):159-161.

90. Hosoya R, Hamana K: **Distribution of two triamines, spermidine and homospermidine, and an aromatic amine, 2-phenylethylamine, within the phylum Bacteroidetes**. *The Journal of General and Applied Microbiology* 2004, **50**(5):255-260.

91. Hamana K, Itoh T, Benno Y, Hayashi H: **Polyamine distribution profiles of new members of the phylum Bacteroidetes**. *The Journal of General and Applied Microbiology* 2008, **54**(4):229-236.

92. Hosoya R, Hamana K, Niitsu M, Itoh T: **Polyamine analysis for chemotaxonomy of thermophilic eubacteria: Polyamine distribution profiles within the orders Aquificales, Thermotogales, Thermodesulfobacteriales, Thermales, Thermoanaerobacteriales, Clostridiales and Bacillales**. *The Journal of General and Applied Microbiology* 2004, **50**(5):271-287.

93. Hamana K, Takeuchi M: **Polyamine profiles as chemotaxonomic markers within alpha, beta, gamma, delta, and epsilon subclasses of class Proteobacteria: distribution of 2-hydroxyputrescine and homospermidine.** *Microbial Culture Collection* 1998, **14**:1-14.

94. Scherer P, Kneifel H: **Distribution of polyamines in methanogenic bacteria**. *Journal of Bacteriology* 1983, **154**(3):1315-1322.

95. Hamana K, Matsuzaki S: **Unusual polyamines in slime molds *Physarum polycephalum* and *Dictyostelium discoideum***. *Journal of Biochemistry* 1984, **95**(4):1105-1110.

96. Ober D, Gibas L, Witte L, Hartmann T: **Evidence for general occurrence of homospermidine in plants and its supposed origin as by-product of deoxyhypusine synthase**. *Phytochemistry* 2003, **62**(3):339-344.

97. Tait GH: **The formation of homospermidine by an enzyme from *Rhodopseudomonas viridis*** *Biochemical Society Transactions* 1979, **7**(1):199-201.

98. Böttcher F, Ober D, Hartmann T: **Biosynthesis of pyrrolizidine alkaloids: putrescine and spermidine are essential substrates of enzymatic homospermidine formation**. *Canadian Journal of Chemistry* 1994, **72**(1):80-85.

99. Yamamoto S, Nagata S, Kusaba K: **Purification and Characterization of Homospermidine Synthase in *Acinetobacter tartarogenes Atcc-31105***. *Journal of Biochemistry* 1993, **114**(1):45-49.

100. Tholl D, Ober D, Martin W, Kellermann J, Hartmann T: **Purification, molecular cloning and expression in *Escherichia coli* of homospermidine synthase from *Rhodopseudomonas viridis***. *European Journal of Biochemistry* 1996, **240**(2):373-379.

101. Hamana K, Matsuzaki S: **Distinct difference in the polyamine compositions of bryophyta and pteridophyta**. *Journal of Biochemistry* 1985, **97**(6):1595-1601.

102. Eisen J: **The Tree of Life**. In: *http://phylogenomicsblogspotcom/2008/10/twisted-tree-of-life-award-2-sciencehtml.* 2008.

103. Forouhar F, Hussain M, Seetharaman J, Fang Y, Janjua H, Xiao R, Cunningham K, Ma L-C, Owens L, Liu J *et al*: **Crystal Structure Of The Homospermidine Synthase Hss From *Legionella Pneumophila* In Complex With Nad, Northeast Structural Genomics Target Lgr54.** In.: Northeast Structural Genomics Consortium 2007.

104. Kaiser A: **Cloning and expression of a cDNA encoding homospermidine synthase from *Senecio vulgaris* (Asteraceae) in *Escherichia coli***. *Plant Journal* 1999, **19**(2):195-201.

105. Ober D, Hartmann T: **Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase**. *PNAS* 1999, **96**(26):14777-14782.

106. Oshima T: **Enigmas of biosyntheses of unusual polyamines in an extreme thermophile, *Thermus thermophilus***. *Plant Physiology and Biochemistry* 2010, **48**(7):521-526.

107. Ober D, Harms R, Witte L, Hartmann T: **Molecular evolution by change of function. Alkaloid-specific homospermidine synthase retained all properties of deoxyhypusine synthase except binding the eIF5A precursor protein**. *Journal of Biological Chemistry* 2003, **278**(15):12805-12812.

108. Fujihara S, Harada Y: **Fast-growing root nodule bacteria produce a novel polyamine, aminobutylhomospermidine**. *Biochemical and Biophysical Research Communications* 1989, **165**(2):659-666.

109. Hamana K, Miyagawa K, Matsuzaki S: **Occurrence of Sym-Homospermidine as the Major Polyamine in Nitrogen-Fixing Cyanobacteria**. *Biochemical and Biophysical Research Communications* 1983, **112**(2):606-613.

110. Higashi K, Terui Y, Suganami A, Tamura Y, Nishimura K, Kashiwagi K, Igarashi K: **Selective structural change by spermidine in the bulged-out region of double-stranded RNA and its effect on RNA function**. *Journal of Biological Chemistry* 2008, **283**(47):32989-32994.

111. Lindemose S, Nielsen PE, Mollegaard NE: **Polyamines preferentially interact with bent adenine tracts in double-stranded DNA**. *Nucleic Acids Research* 2005, **33**(6):1790-1803.

Chapter Four

Cloning, Characterisation & Comparison of the Enzyme Homospermidine Synthase

TABLE OF TABLES

TABLE OF FIGURES

Chapter Four - Cloning, Characterisation & Comparison of the Enzyme Homospermidine Synthase

4.1 Background

There have been numerous extensive reviews on the common polyamines (putrescine, spermidine and spermine) and their metabolism [1-5]. However, relatively little is known about the symmetrical triamine homospermidine. It is of particular interest because many phylogenetically diverse organisms have been identified as homospermidine-containing [6-10].

A homospermidine synthase (HSS) enzymatic activity was first reported by Tait [11] in the α-Protobacterium *Rhodopseudomonas viridis*. Homospermidine was formed when the cell-free extracts were incubated in the presence of putrescine and NAD^+. It is proposed that two molecules of putrescine join to make one homospermidine [12] (Figure 3.6).

The gene encoding the HSS enzyme of *R. viridis* has since been cloned, and the encoded enzyme was shown to be a homodimer with subunits of 52 kDa [13]. Although HSS activity was shown *in vivo* (*Escherichia coli* expression), HSS activity was not demonstrated *in vitro*.

Kinetic analysis of HSS has been reported from purified extracts of the γ-proteobacterium *Acinetobacter tartarogenes* [14] and α-Protobacterium *Rhodospeudomonas viridis* [12], as discussed in Chapter 3, Table 3.2. However, there is no record of characterisation and experimentation on the recombinantly expressed and purified enzyme.

Homospermidine has been found in a wide range of diverse species, this suggests that they contain a HSS protein or similar enzyme to synthesise it. It is necessary to study the function of the HSS enzyme, and to research its distribution over a wide range of species as this will allow us to understand more about the evolution and divergence between bacterial species. Pathogens, e.g. *Legionella* spp., and environmentally relevant bacteria, e.g. *Rhizobia* spp., are known to contain *hss* orthologues. If more is understood about the genetic

and biochemical composition of these *hss*-containing organisms, areas such as drug development and resistance can be investigated. After suspected *hss* orthologues have been identified, it is essential that a selection of these be tested biologically for HSS activity to back up the bioinformatic analysis with biochemical tests.

Bottcher *et al.* [12] demonstrated that the bacterial homospermidine synthase (in *R. viridis* cell extract) was able to produce alternative polyamine products, such as diaminopropane, when supplied with artificial substrates in the enzyme assay. This suggested non-specific substrate binding, and if HSS is able to make other products, it would be interesting to learn if bacterial species can make alternative polyamine products under certain conditions.

It is essential to understand more about HSS, and other biosynthetic enzymes, because the implications and applications of this research are wide. As our knowledge about unusual, previously unknown or unculturable organisms increases, so does our ability to use this information for alternative applications and interpretations of scientific research. This area of research is essential in the advancement of understanding the polyamine biosynthetic pathway, and its evolutionary development across species.

4.2 Objectives

- To recombinantly express and purify the *Bradyrhizobium japonicum* homospermidine synthase (HSS) enzyme in an *Escherichia coli* host.

- To kinetically characterise purified *B. japonicum* HSS enzyme.

- To confirm the function of predicted *hss* genes from phylogenetically diverse organisms by heterologous expression in an *E. coli* host.

4.3 Results

4.3.1 Enzymatic characterisation of *Bradyrhizobium japonicum* homospermidine synthase (HSS)

The *hss* from *Bradyrhizobium japonicum,* a nitrogen-fixing root nodule symbiont, was chosen because homospermidine has been detected in cell-free extracts from this species [15]. Figure 4.1 shows the polyamine profile of *B. japonicum* at stationary phase, and for comparison, another α-Proteobacterium *Rhizobium leguminosarum*, both grown in defined minimal AMS medium. Rich medium, such as TY, contains substantial amounts of putrescine and spermidine, which can obscure the chromatographic data. Both species were shown to contain homospermidine as their only triamine, i.e. no spermidine or norspermidine was detected, and each contained the homospermidine precursor, putrescine. This result agreed with earlier studies [10].

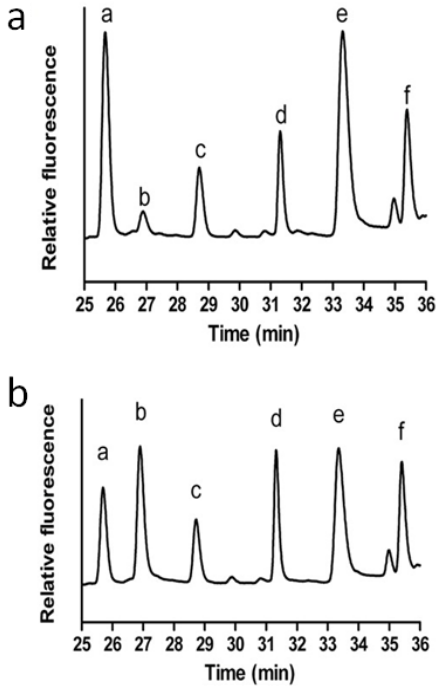


Figure 4.1 HPLC analysis of intracellular polyamines
(a) HPLC analysis of polyamines in cell extracts of *B. japonicum* grown to early stationary phase (OD_{600nm} = 0.4; 9 days) in AMS minimal defined medium. (b) HPLC analysis of cell extracts of *R. leguminosarum*, grown to stationary phase (OD_{600nm} = 0.8; 4 days) in AMS defined minimal medium. Identification of polyamines using HPLC; *a*, diaminopropane; *b*, putrescine; *c*, unknown peak; *d*, fluorescent label; *e*, homospermidine; *f*, diaminoheptane (internal standard).

The *hss* orthologue from *B. japonicum* was identified within the published genome (GenBank: BA000040.2) based on sequence similarity to the previously characterised *hss* genes [12, 13].

As a first step in an evolutionary analysis of HSS, the *hss* orthologue of the α-Proteobacterium *B. japonicum* (Accession Number NP_774402) was recombinantly expressed in *Escherichia coli* (confirmed by HPLC analysis) and the kinetic behaviour of the recombinant enzyme was compared with that of HSS activity reported for purified extracts from *A. tartarogenes* and *R. viridis.*

The *B. japonicum hss* orthologue was amplified from genomic DNA and ligated into the pET21a expression vector. The *B. japonicum hss* gene was induced and over-expressed in an *E. coli* host. The predicted weight of the protein was 53.3 KDa. Figure 4.2 shows the purified HSS enzyme (~55 KDa) on an SDS Nu-page gel, and a HPLC chromatograph of the HSS assay demonstrating homospermidine production which was absent in the control experiment (the assay without the enzyme). This result concluded that the *B. japonicum hss* definitely has HSS activity. Purified recombinant N-terminally-T7-tagged protein was assayed with putrescine as substrate, and the product homospermidine was quantified by fluorescence HPLC.

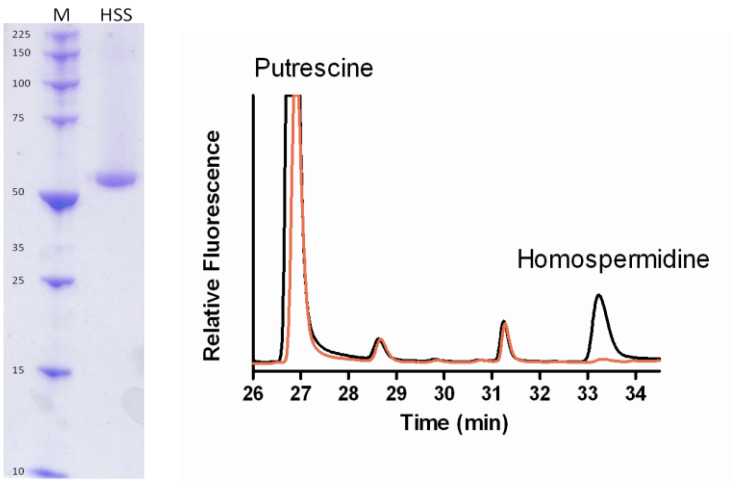


Figure 4.2 HSS protein and assay
(Left) The SDS-page shows the recombinant T7-purfied HSS protein (HSS) used in the HSS assays and M=marker Promega, Broad Range , KDa. (Right) HPLC chromatograph showing the production of homospermidine in the assay (black) compared to the negative control (red - no enzyme).

Figure 4.3 shows the kinetic parameter determination experimental procedures performed to investigate the optimum conditions of the *B. japonicum* HSS enzyme. The amount of homospermidine produced during the enzyme assays was quantified using a homospermidine standard curve. For this, known concentrations of homospermidine (μM) were injected into the HPLC in triplicate and the average peak area (mAU) was used to generate a curve. The detection of homospermidine was monitored. The relationship was linear between 0.125 and 5 μM (Figure 4.3a), so the quantification of homospermidine was accurate.

To ensure that the enzyme reaction was progressing in a linear nature, and that an appropriate incubation time was chosen for further experiments, the amount of homospermidine formed after incubation at four different time points was investigated. Figure 4.3b shows that within the first 20 minutes of the assay, the reaction is linear with respect to time. For the enzyme characterisation experiments, an incubation time of four minutes was used. For this investigation, dilutions of enzyme (in an appropriate buffer) were tested and the amount of homospermidine produced was graphed against the amount of

enzyme. The reaction was shown to be linear with respect to protein concentration (Figure 4.3c).

The optimum pH of the HSS enzyme was tested using a selection of Tris buffers over a range of pH. At the optimum pH, the HSS activity is described as 100% and at the other pHs the activity is shown as a percentage of this. The optimum *in vitro* pH for *B. japonicum* HSS was shown to be 8.4 (Figure 4.3d).

The kinetic parameters of the enzyme were calculated according to the Michaelis-Menten equation (Figure 4.3e) and are summarised in Table 4.1. The basis of this experiment is to test the enzyme under optimum conditions, with increasing substrate concentration. The data was plotted on a graph; substrate concentration (putrescine [PUT] μM) against enzyme activity (μmoles HSPD $minute^{-1}$ mg HSS $protein^{-1}$), and the Michaelis–Menten kinetic model was used to describe the enzyme kinetics. This kinetic model is only relevant to situations where very simple kinetics can be assumed, such as no product/substrate inhibition or allostericity [16]. For completeness, a Hanes-Woolf plot of the data is shown (Figure 4.3f). The Hanes-Woolf plot can be used to calculate the kinetic parameters of an enzyme. The relation between substrate concentration and substrate concentration divided by the velocity of the reaction is linear. Specific points on the line of this linear relationship correspond to the various kinetic parameters.

The enzymatic characterisation of *B. japonicum* HSS and results in this section have been reported [17].

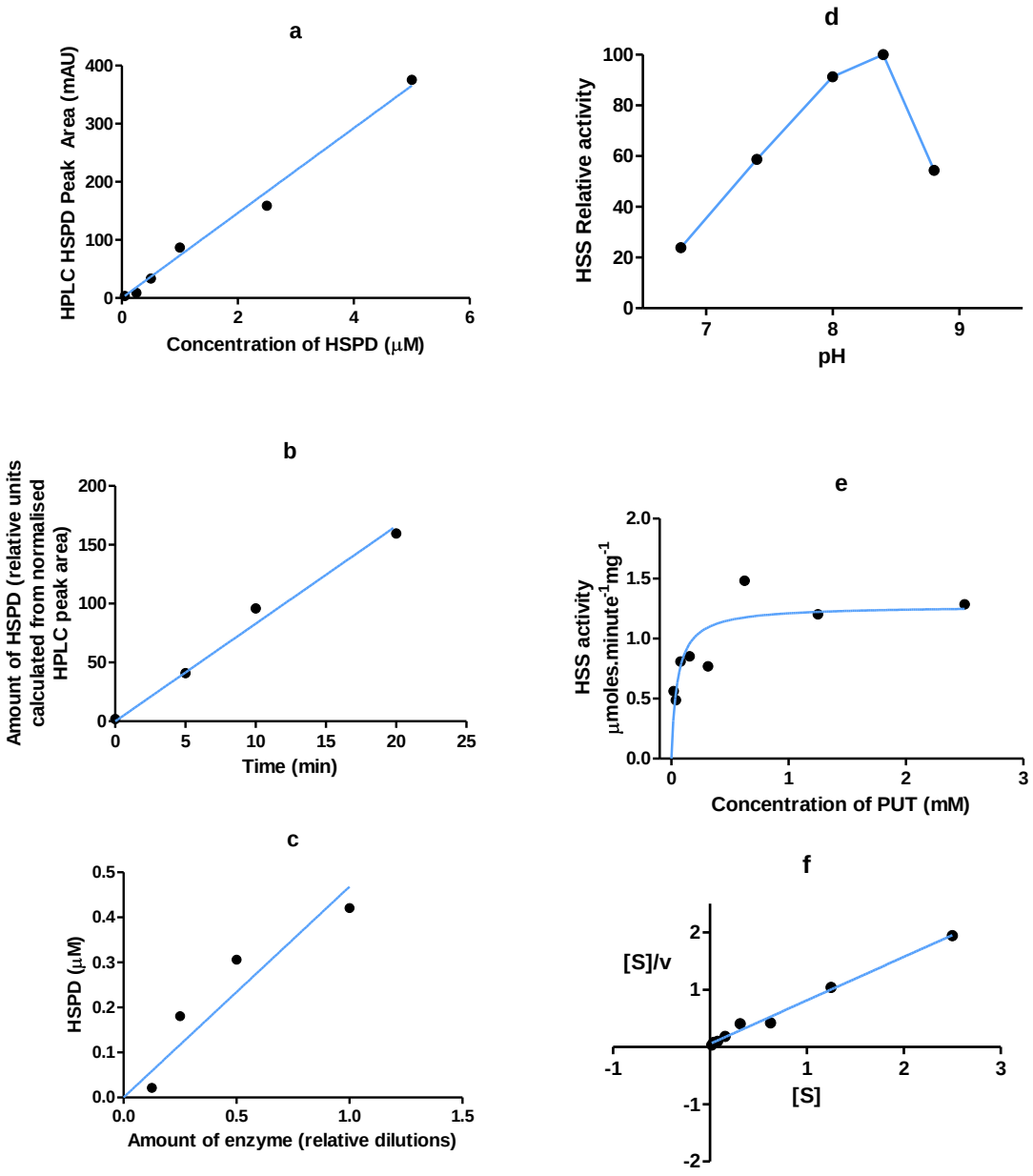


Figure 4.3 Kinetic parameters of HSS
a). An example of a homospermidine (HSPD) standard curve, b). The linear relationship between the length of the HSS assay, in minutes, and the amount of HSPD produced, based on peak area relative to the amount of internal standard, c). The linear nature between the concentration of HSS enzyme and the μM HSPD produced. The concentration of HSPD was calculated by extrapolation of a HSPD standard curve, d). The activity of the enzyme based on the amount of HSPD produced at a range of pH, e). The application of the Michaelis-Menton curve to these data (PUT = putrescine). f.) The data shown using a Hanes-Woolf plot [S] – substrate concentration, v – reaction velocity.

According to Michaelis and Menten the conversion of a substrate (S) into a product (P) is divided into two steps as shown:

$$E + S \underset{}{\overset{K_m}{<=>}} ES \overset{k_{cat}}{<=>} E + P$$

The first reaction step describes the binding of the substrate (S = putrescine) to the enzyme (E = HSS) and the constant K_m corresponds to the dissociation constant of the equilibrium under conditions where the product formation is very slow compared to the dissociation process of the substrate. In addition, K_m equals the substrate concentration at half maximal reaction rate $V_{max}/2$. In other words, K_m describes the affinity of the substrate for the enzyme.

The second reaction step describes the catalytic rate (also known as the rate of product formation) and referred to as the turnover number, k_{cat}. The turnover rate is defined as the maximal number of product (P = homospermidine) per active site per unit time.

The Michaelis-Menten kinetic principle is valid only under saturation conditions, i.e., when the concentration of substrate is much larger than the enzyme concentration. This is why it was necessary to define the parameters of the reaction before kinetic interpretation. The maximum reaction rate V_{max} describes a steady-state equilibrium of the reaction catalyzed by the enzyme. The steady-state equilibrium is an important concept in biochemistry because many enzyme-catalysed reactions run at saturation and the product often is removed from the reaction site so as to render the reaction irreversible.

The ratio k_{cat} / K_m defines a measure of the catalytic efficiency of an enzyme-substrate pair. This is known as the substrate specificity. It refers to the properties and reactions of free enzyme and free substrate. This is the value used to directly compare one enzyme with another in this investigation. The theoretical limit for k_{cat} / K_m is set by the rate constant of the initial complex formation (ES) and cannot be faster than the diffusion controlled interaction of substrate and enzyme. The specificity of an enzyme is therefore a measure

of the specificity of an enzyme for competing substrates or of competing enzymes for a single substrate.

Table 4.1 compares the kinetic behaviour of the recombinant *B. japonicum* HSS with published HSS activity [12, 14] from purified extracts of *A. tartarogenes* and *R. viridis*. The overall catalytic efficiencies (k_{cat}/K_m) of the *B. japonicum* recombinant HSS and the endogenous *A. tartarogenes* HSS are ~5-fold higher than the *R. viridis* enzyme. This is the result of a lower K_m in the case of the *B. japonicum* enzyme and a higher k_{cat} in the case of the *A. tartarogenes* enzyme.

As well as the production of homospermidine, HSS also produces ammonia. An additional enzyme assay to detect ammonia liberation was performed to confirm the reaction (as shown in Figure 3.6). Results showed that ammonia was produced (Figure 4.4), however the assay was not sensitive or robust enough, in this circumstance, to be used for enzymatic characterisation or statistical interpretation. The presence of ammonia confirmed the predicted reaction to be true.

	V_{max}	k_{cat}	K_m for putrescine	k_{cat}/K_m	Optimum pH
	µmol.min^{-1}mg^{-1}	*s^{-1}*	*mM*	*M^{-1}s^{-1}*	
B. japonicum recombinant HSS	1.4 ± 0.13	1.18 ± 0.12	0.05 ± 0.02	23,600 ± 12,000	8.4
R. viridis HSS	1.32	1.2	0.26	4,600	8.8
A. tartarogenes HSS	8.8	8	0.28	28,600	8.7

Table 4.1 Comparison of kinetic parameters of homospermidine synthase activities.
For the *B. japonicum* recombinant HSS enzyme, homospermidine product formation was monitored by HPLC. A minimum of three assays was performed for each enzyme preparation. The kinetic data for *R. viridis* and *A. tartarogenes* are based upon experiments performed from cell extracts.

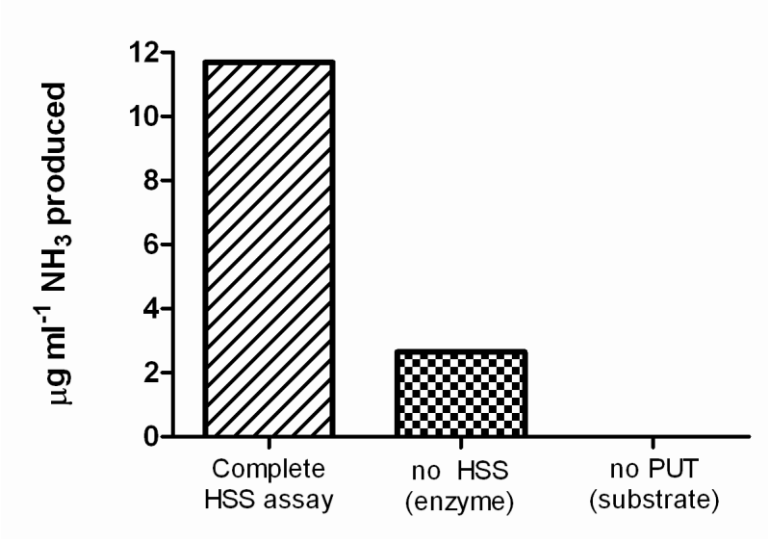


Figure 4.4 The amount of ammonia (NH_3) produced during the HSS assay.
Incomplete assays with no enzyme (no HSS) and no substrate (no PUT) were also monitored. The assay without substrate gave the lowest reading and was used as a 'blank' to calculate the amount of ammonia in the other samples using an equation derived by the manufacturers.

4.3.2 The expression of phylogenetically distant *hss* orthologues in *E. coli* results in homospermidine production.

Having validated the enzymatic activity of the recombinatly expressed and purified *B. japonicum* HSS, the amino acid sequence of this protein was used to search available genomes for HSS orthologues using PBLAST. In addition, expressed sequence tags (ESTs) were searched by TBLAST and a neighbour-joining phylogenetic tree of *hss* orthologues created (Figure 4.5 - the alignment on which the tree is based is shown in supplementary information Appendix I). The most notable aspect of the phylogenetic distribution of HSS is that it is mainly represented in the α-Proteobacteria. Seventeen representative α-proteobacterial *hss* orthologues are depicted in Figure 4.5, but in total, there were *hss* genes in 69 sequenced α-Proteobacterial genomes which align closely with one another (Appendix II).

As little as 26% sequence identity is shared by distant *hss* orthologues, so it is important that credible biochemical proof is provided of their corresponding HSS enzymatic activity. It is a collective fear that in this era of increasingly rapid genome sequence and annotation, that genes may be accidentally mis-annotated due to a rather low sequence similarity, or tenuous linkage to a particular genome. Five of the *hss* orthologues, represented in Figure 4.5 shown with four red asterisks, along with the confirmed *B. japonicum hss*, were expressed in *E. coli*, a species that does not naturally synthesise homospermidine. The additional five *hss* genes chosen were *Opitutus terrae* [18], an obligate bacterial anaerobe from rice paddy fields and belonging to the diverse Verrucomicrobia phylum, a gene from an unknown species identified from the GOS (Global Ocean Sampling) marine metagenome, Ralstonia phage φRSL1, the eukaryotic amitochondriate heterolobosean amoeba *Sawyeria marylandensis* [19] and the eukaryotic, single cell, ciliate *Paramecium tetraurelia* [20]. Except for the *B. japonicum hss* orthologue, which was cloned by PCR from genomic DNA, all other *hss* genes were synthesised with *E. coli*-optimised codons. Gene synthesis was essential in the case of the *P. tetraurelia hss* gene because ciliates use a different genetic code, and some codons which form amino acids in *P. tetraurelia,* would be

interpreted as 'STOP' codons in *E. coli*. And also in the case of the GOS marine metagenome sequence, because its genomic origin is unknown.

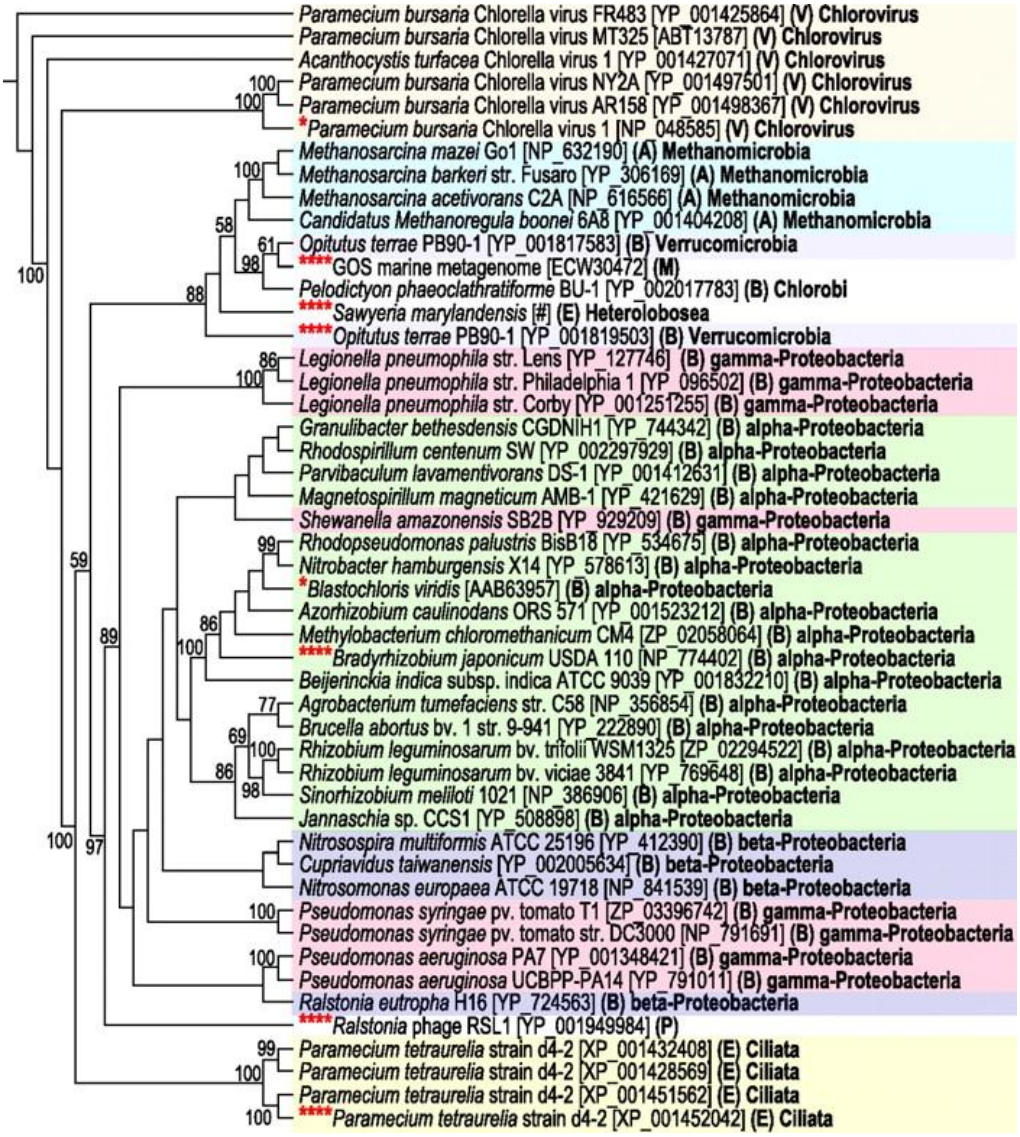


Figure 4.5 Neighbour-joining tree of HSS orthologues.
Analysis performed by Tony Michael.
The unrooted tree was built using PAUP* (as described by Lee [21]) with percentage values from 1000 bootstrap replicates indicated; values less than 50% are not shown. The alignment on which the tree is based is presented in Appendix I. Only representative α-Proteobacteria are shown (a more complete representation is shown in Appendix II). Orthologues with four red asterisks were characterised in this study, and those with one asterisk were characterised previously.

NOTE: *Rhodospeudomonas viridis* (as referred to in this thesis) is the same organism as *Blastochloris viridis* (as seen in the diagram highlighted with an asterisk).

The six *hss* genes were expressed in *E. coli* and HPLC of the cell extracts confirmed which polyamines were present. The host *E. coli* cells without the recombinant HSS accumulate putrescine, cadaverine, and spermidine (Figure 4.6 black peaks), and expression of each of the six *hss* orthologues in *E. coli* resulted in accumulation of homospermidine in the host cells (Figure 4.6 peak f, confirmed by LC-MS/MS, Figure 4.7). As Figure 4.6 shows, different amounts of homospermidine appear in the various samples; this could be due to the different levels of protein expression by the *E. coli* host, or by the affinity and reactivity of the various HSS with the substrate, putrescine.

Some of the *hss* orthologues also resulted in an accumulation of diaminopropane (Figure 4.6 peak a: *B. japonicum*, Ralstonia phage φRSL1, and *O. terrae*). The presence of diaminopropane suggests that the HSS enzyme is using spermidine, a polyamine found naturally in *E. coli*, as a substrate to produce homospermidine. When this happens, diaminopropane is released as an additional product. The mechanism of this reaction is discussed later in this chapter. Due to the products seen by HPLC analysis, it was concluded that the HSS enzymes from *B. japonicum*, Ralstonia phage φRSL1 and *O. terrae* were all able to use spermidine to produce homospermidine, whereas *P. tetraurelia,* the GOS gene and *S. marylandensis* were all unable to do so. This difference in substrate specificity between the enzymes was particularly interesting, as it suggests an adaptation of the enzyme which may have occurred randomly during evolution based on the distribution of the organisms in the phylogentic tree (Figure 4.5). The ability of HSS to use spermidine to make homospermidine may be dependent on the polyamines available. So, it is unknown whether the phenomenon is occurring naturally in the species that the *hss* were taken from. However, further testing of the six enzymes *in vitro* would be appropriate to make conclusions on the substrate affinity of each *hss* orthologue. This is the first demonstration *in vivo* of the production of diaminopropane from spermidine by HSS.

All of the six hetrologously expressed enzymes in *E. coli*, except *P. tetraurelia,* produced an additional peak of unknown identity (Figure 4.6 peak g). The unknown peak was analysed by LC-MS/MS and was determined to be aminobutylcadaverine (Figure 4.7). Aminobutylcadaverine is produced by

HSS when either a putrescine and cadaverine, or spermidine and cadaverine, are used as substrates. The mechanism of this reaction is discussed later in this chapter. The reaction is able to take place because cadaverine, spermidine and putrescine are all found naturally in the *E. coli* host. This is the first demonstration of the heterologous biosynthesis of aminobutylcadaverine *in vivo*. Cells of *B. japonicum* do not naturally synthesise cadaverine, so the HSS co-product aminobutylcadaverine is not detected in those cells (Figure 4.1).

The different amounts of products seen in Figure 4.6 may be due to the different levels of protein expression in the samples, due to the affinity for the HSS enzyme to the available products or the kinetic parameters of the HSS during *E. coli* expression.

The heterologous expression of the six phylogenetically diverse HSS genes in *E. coli* has been reported [17] and the paper also speculates, as in this report, about the possible routes of genetic transfer by which the species have gained their *hss*.

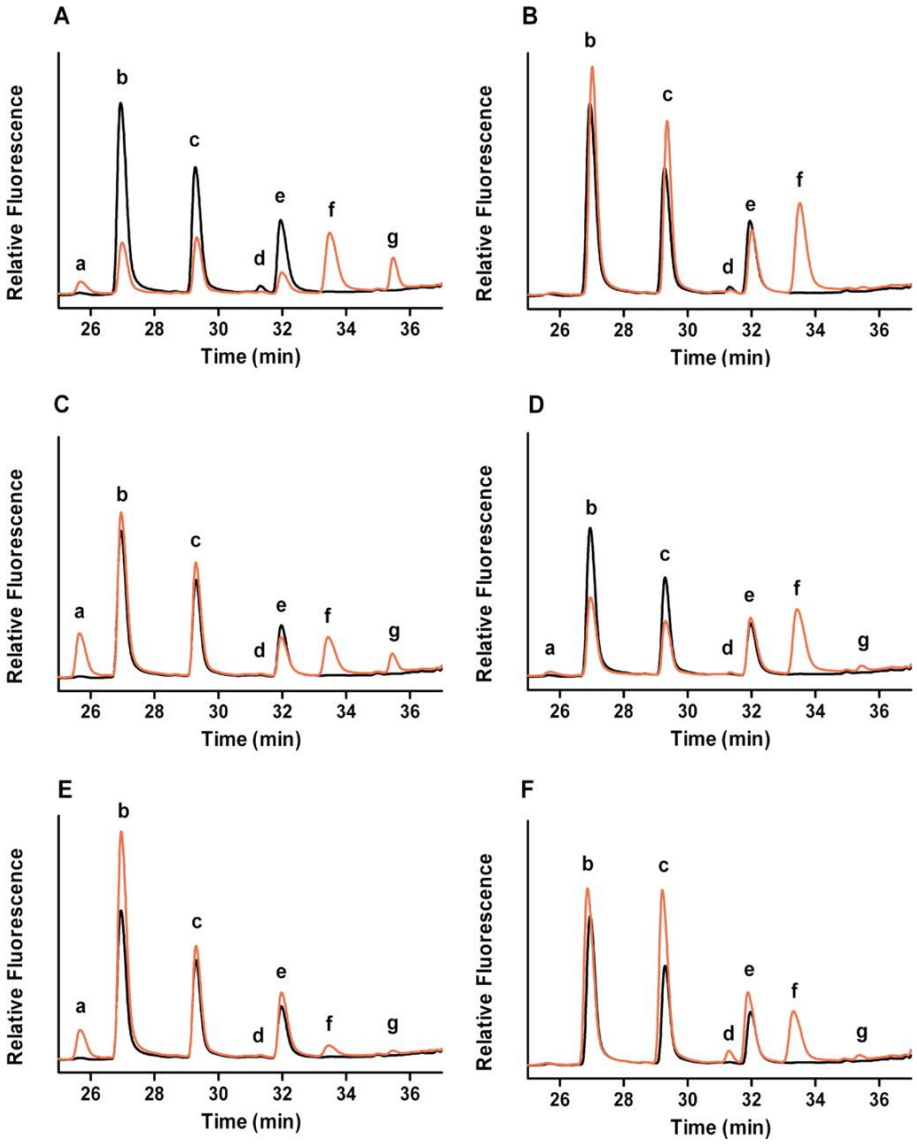


Figure 4.6 HPLC analysis of cell extracts from *E. coli* expressing recombinant HSS orthologues.
Analysis performed with Kath Elliott.
In each case, the extract from cells expressing the induced homospermidine synthase orthologue is shown in red, and the empty vector control extract is in black. *A*, *B. japonicum*/pET21a; *B*, *P. tetraurelia*/pET21a; *C*, *Ralstonia* phage φRSL1/pET21b; *D*, *S. marylandensis*/pET21b; *E*, *O. terrae*/pET21b; *F*, GOS marine metagenome/pET21b. *a*, diaminopropane; *b*, putrescine; *c*, cadaverine; *d*, fluorescent label; *e*, spermidine; *f*, homospermidine; *g*, aminobutylcadaverine.

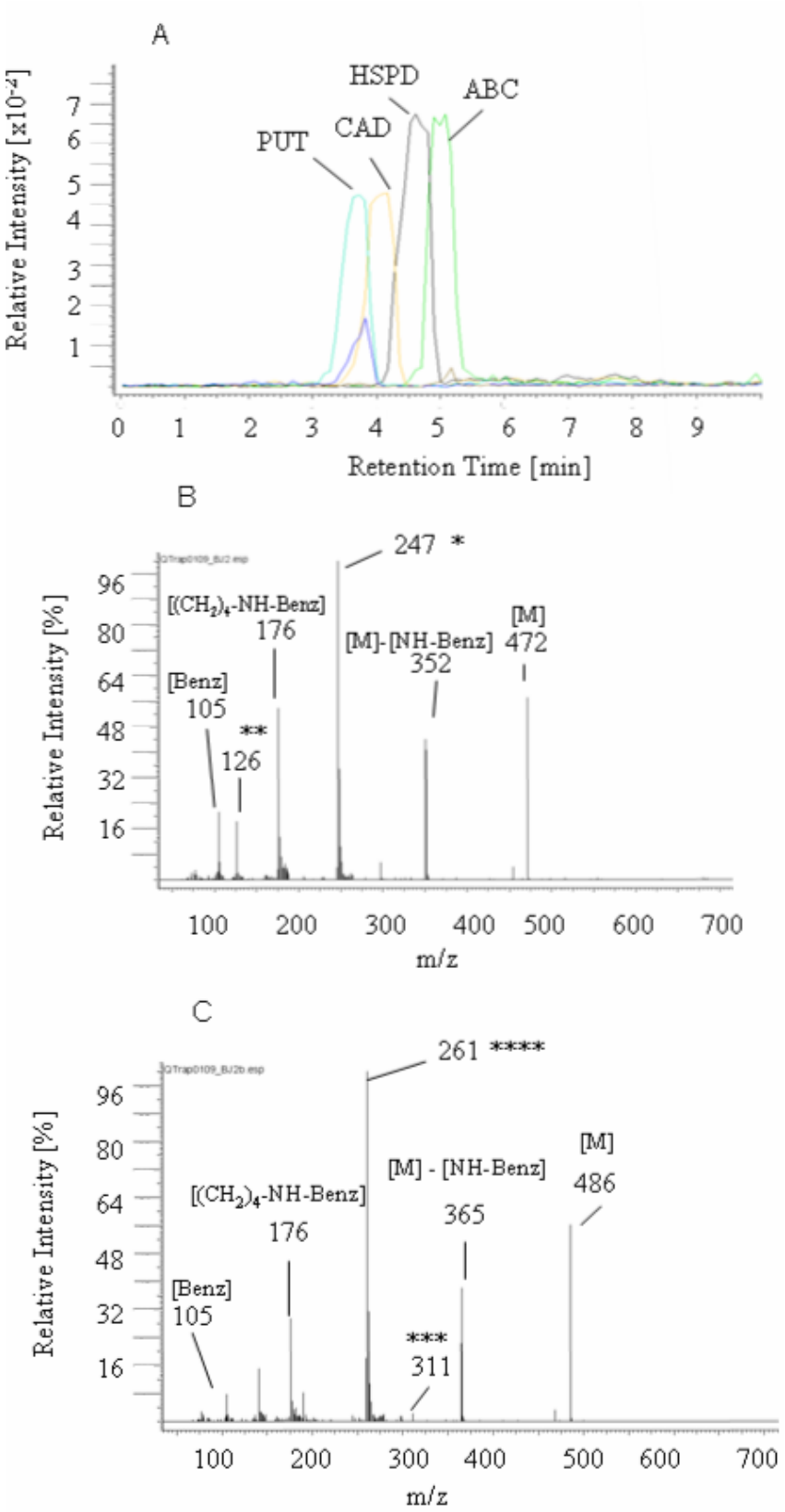


Figure 4.7 Mass spectrometric confirmation of homospermidine and *N*1-aminobutylcadaverine identities.
Analysis performed by Christine Fuell.
Top panel (A): Mass traces from LC-MS/MS analysis (Q-Trap) of *E. coli* cell extracts expressing the recombinant *B. japonicum* homospermidine synthase.

Polyamines were derivatized by benzoylation for detection by UV during LC separation step. Abbreviations: benz, benzoyl group; PUT, putrescine; CAD, cadaverine; HSPD, *sym*-homospermidine; ABC, *N*1-aminobutylcadaverine. Colour code: cyan, m/z = 296 (PUT x 2 benz); yellow: m/z = 310 (CAD x 2 benz); black: m/z = 472 (HSPD x 3 benz); green: m/z = 486 (ABC x 3 benz). Middle panel (B): MS2 of *sym*-homospermidine (HSPD) from the *E. coli* cell extracts expressing *B. japonicum* recombinant HSS. Comparison to a pure standard was used to confirm the fragmentation pattern for additional validation. In the first dimension MS (not shown) of the HSPD the peaks of m/z = 472 [M] and m/z = 494 [M+Na+] were present. The daughter spectrum of m/z = 472 is shown here. Fragments are assigned to their respective mass peaks. *: m/z = 247 corresponds to a diaminobutyl x 2 benz residue, of which one of the benzoyl groups has been partly eliminated after rearrangement. **: m/z = 126 corresponds to an aminobutyl x 1 benz residue of which one of the benzoyl groups has been partly eliminated after rearrangement. Bottom panel (C): MS2 of *N*1-aminobutylcadaverine (ABC) corresponding to the unknown peak in the *E. coli* cell extracts expressing the *B. japonicum* HSS. In the first dimension MS (not shown) of the ABC the peaks of m/z = 486 [M] and m/z = 508 [M+Na+] were present. The daughter spectrum of m/z = 486 is shown here. The fragments are assigned to their respective mass peaks. ***: m/z = 311 corresponding to cadaverine x 2 benz ([M] - [aminobutyl x 1 benz]). ****: m/z = 261, corresponds to a cadaverine x 2 benz residue of which one of the benzoyl groups has been partly eliminated after rearrangement.

4.4 Discussion

4.4.1 Characterisation of Bacterial HSS from the α-Proteobacterium *Bradyrhizobium japonicum*

This is the first time that the enzymatic activity of a recombinantly expressed and purified *B. japonicum* HSS has been established, all previous studies on HSS have been undertaken using cell extracts. By using two different assay techniques; firstly one to detect homospermidine production and secondly to detect ammonia release, the predicted reaction has been confirmed (Figure 3.6). Based on the results from this investigation, it can be hypothesised that two molecules of putrescine were joined together by the HSS to form one homospermidine molecule with the release of ammonia.

The optimum pH of *B. japonicum* HSS was established to be 8.4 and the necessary conditions for calculating the kinetic parameters were ascertained by ensuring the reaction was progressing in a linear fashion. Conditions were chosen so that the reaction was not limited by time nor enzyme concentration. The HSS was characterised and the catalytic efficiencies were compared with those from previously characterised HSS from alternative species. The optimum pH of 8.4 is most likely not the intracellular pH of *B. japonicum*, but is optimum for *in vitro* acivity.

Comparison of the catalytic efficiencies of the *B. japonicum* recombinant HSS with the previously characterised endogenous *A. tartarogenes* HSS showed that these are similar. The k_{cat}/K_m of *B. japonicum* HSS was 23,600 ± 12,000 $M^{-1}s^{-1}$ and the published result from *A. tartarogenes* HSS was 28,600 $M^{-1}s^{-1}$. This result was largely unsurprising, because both organisms are Gram-negative, soil dwelling bacteria, although it was predicted that the recombinant enzyme would be more efficient at producing homospermidine than an enzyme remaining within cytoplasmic solution.

The *B. japonicum* (and *A. tartarogenes*) HSS had catalytic efficiencies ~5-fold higher than the *R. viridis* enzyme. The catalytic efficiency is the ratio of k_{cat}/K_m and used as a defined measure of the catalytic efficiency of an enzyme-substrate pair. The difference in catalytic efficiency can be explained by the

lower K_m in the case of the *B. japonicum* enzyme and a higher k_{cat} in the case of the *A. tartarogenes* enzyme.

It cannot be concluded whether the catalytic parameters of purified recombinant *B. japonicum* HSS are representative of the natural HSS activity in the *B. japonicum* cells. Obviously, the cloning, expression and purification of the *B. japonicum* HSS may have inadvertently decreased the enzyme activity. In particular, the T7 column uses an acidic buffer (pH 2.2) to elute the protein from the column. The freezing and long-term enzyme storage was performed by flash-freezing (liquid nitrogen) in a stable buffer (20% glycerol), in an attempt to preserve and protect the enzyme structure, however the activity of the HSS was not tested before and after a single freeze-cycle. This step may have adversely affected the catalytic values generated during the kinetic calculations.

4.4.2 Confirmation of HSS activity of six phylogenetically diverse *hss*-annotated orthologues and gene inheritance patterns

The next goal of this investigation was to compare HSS enzymes between different species. Initially, it was important to confirm that all six *hss*-annotated genes functioned with HSS activity. The ability of a further five phylogenetically diverse *hss* orthologues to synthesise homospermidine when expressed in an *E. coli* host has been demonstrated (Figure 4.6). These five genes were either annotated as putative homospermidine synthases, or found based on sequence similarity to other annotated or characterised HSS genes.

Vertical gene inheritance: This result indicates that *hss* has a widespread distribution and suggests the possibilities of further studies on the phylogenetic distribution of this polyamine-synthetic gene. Looking at the phylogenetic distribution of the *hss* orthologues in Figure 4.5, certain conclusions about the evolution of the *hss* can be made. For example, in the α-Proteobacteria there is extensive vertical inheritance of *hss*. This can be concluded by the presence of an *hss* gene in many of the genomes examined in this division. A more complete representation of the α-proteobacterial *hss* orthologues is shown in the neighbour-joining tree (Appendix II).

There are no *hss* orthologues in the most basal α-proteobacterial order, the Rickettsiales, which are mainly represented by species with reduced genomes. Most Rickettsiales survive only as endosymbionts and can be difficult to cultivate in the laboratory. There are also no *hss* orthologues in the Sphingomonadales and Caulobacterales orders, which diverged after the *hss*-containing Rhodospirillales [22, 23]. This suggests that the *hss* gene was lost in the common ancestor of those orders.

Horizontal gene transfer: Outside of the α-Proteobacteria, the distribution of *hss* orthologues in bacteria is sporadic. There are a few β- and γ-Proteobacterial species containing *hss* orthologues and only one species contain *hss* in the Verrucomicrobia (*O. terrae*) and also in the Chlorobi (*Pelodictyon phaeoclathratiforme*). There are no reports of *hss* orthologues in the ε-Proteobacteria. Regarding the sporadic *hss* distribution, it can be concluded that the *hss* orthologues have moved due to horizontal gene transfer between species. Four closely related methanomicrobial species of the archaea contain *hss*, this information is consistent with the known presence of homospermidine in methanogens [24] and high rates of horizontal gene transfer between these species [25].

Studying *hss* distribution has highlighted an interesting situation regarding gene transfer. The evolutionary origin of the chloroviral *hss* is uncertain. Chloroviral genomes are from viruses that infect the *Chlorella* endosymbiont of the bacteriovorus single cell eukaryote *Paramecium bursaria*. The algal host *Chlorella* seems an unlikely source because no plant or algal species is currently known to possess an *hss* orthologue. In contrast, plants contain the deoxyhypusine synthase-like-homospermidine synthase (DHS-HSS) as discussed in Chapter 3. The genome of *Paramecium tetraurelia* contains four paralogous copies of *hss*. If the *Paramecium* cell was the source of the *hss* gene in the chlorovirus genome, the genetic code of the gene would have to be reprogrammed. Ciliates such as *Paramecium* use a different genetic code where the TAA and TAG stop codons encode glutamine, whereas chloroviruses use the 'normal' genetic code. So, it is a mystery as to exactly what happened, but it would be interesting to speculate that the gene could have crossed

between two, physically speaking, closely located organisms, and had the ability to mutate the genetic code, in order to express a working HSS.

The complete *hss* orthologue from the GOS marine metagenome (selected for hetrologous expression in *E. coli*), possesses 69% identity with the closest match from sequenced genomes: an *hss* from the verrucomicrobium *Opitutus terrae* (another one of the six genes chosen for hetrologous expression in *E. coli*). Extrapolation of this information suggests that the GOS micro-organism may be similar to *O. terrae* (based on the *hss* gene similarity), or may have derived from a shared common ancestor. Alternatively, the GOS micro-organism may have acquired the *hss* by horizontal gene transfer. There were several *hss* orthologues identified in the GOS metagenomic analysis. All of the putatively full-length *hss* orthologous sequences in the GOS marine metagenome exhibit at least 50% identity to this *O. terrae* sequence. It is notable that the sequenced genomes of *Pelagibacter ubique*, *Prochlorococcus* species, and *Synechococcus* species do not contain *hss* orthologues, and ~50% of the GOS marine metagenome sequence that can be recruited to known genomes belongs to these three groups [26].

Application of comparative genomics: Some key bacterial pathogens possess *hss* orthologues including *Legionella pneumophila*, which was shown previously to contain homospermidine [27] and some of the *Brucella* species (causative agent of brucellosis), *Pseudomonas aeruginosa*, and the plant pathogen *Pseudomonas syringae*. The application of comparative genomics may lead to discovery of previously unknown drug-sensitive pathways. Other applications of comparative genomics include interpretation of data to further understand how the organisms came to contain the genes they do either by vertical or horizontal gene transfer. In addition, predictions can be made about the evolution and relationship between different species and the orders which they belong to.

4.4.3 Production of Non-native *E. coli* Polyamine Products

HPLC and LC-MS/MS confirmed the presence of polyamines in *E. coli* cells hetrologously expressing phylogenetically diverse *hss* (Figures 4.6, 4.7 and

Table 4.2). Each complemented *E. coli* strain was able to grow as normal and to synthesise homospermidine. The naturally occurring *E. coli* polyamines: putrescine, cadaverine and spermidine were observed in all. By examining the polyamines present, conclusions can be made about the ability of the HSS enzymes to use various substrates to synthesise homospermidine, or an alternative product, aminobutylcadaverine. When HSS uses spermidine to make homospermidine, the by-product released is diaminopropane. Based on this information, the results suggested that the HSS enzymes from *B. japonicum*, Ralstonia phage φRSL1 and *O. terrae* are all able to use spermidine to produce homospermidine, whereas *P. tetraurelia,* the GOS gene and *S. marylandensis* could not. In addition to this, it can be concluded that all six of the genes, except *P. tetraurelia,* can use cadaverine to make the alternative triamine aminobutylcadaverine.

	DAP	PUT	CAD	SPD	HSPD	ABC
B. japonicum	x	x	x	x	x	x
P. tetraurelia		x	x	x	x	
Ralstonia phage	x	x	x	x	x	x
S. marylandensis		x	x	x	x	x
O. terrae	x	x	x	x	x	x
GOS marine metagenome		x	x	x	x	x

Table 4.2 The polyamines observed during HPLC (and in some cases LC-MS/MS) of *E. coli* extracts hetrologously expressing *hss*.
Refer to Figure 4.6. DAP – diaminopropane, PUT- putrescine, CAD – cadaverine, SPD – spermidine, HSPD – homospermidine, ABC – aminobutylcadaverine.

Böttcher *et al.* [12] proposed a two-site mechanism for homospermidine formation by bacterial HSS. They described the production of homospermidine from two polyamines. The bacterial HSS is able to use either putrescine-putrescine or spermidine-putrescine to produce homospermidine. Because of this substrate promiscuity, the bacterial HSS enzyme is able to synthesise a range of products. If two putrescine molecules are fused to create homospermidine, ammonia is released. When spermidine and putrescine are used as substrates for homospermidine, the by-product released is diaminopropane (Figure 4.8). If cadaverine is available, the HSS is able to

ultilise it along with putrescine or spermidine to form aminobutylcadaverine. If cadaverine is metabolised by HSS, then homospermidine is not formed. Aminobutylcadaverine is produced with by-products diaminopropane or ammonia when either spermidine or putrescine (along with cadaverine) are used as the substrates, respectively. This substrate variation and product synthesis of HSS is described diagrammatically in Figure 4.8.

The substrate putrescine (or spermidine) binds "site A," where it is then oxidatively deaminated by an NAD^+-dependent mechanism. Ammonia is released with putrescine as the substrate, or diaminopropane is released with spermidine as substrate. The resulting semialdehyde reacts with a second amine (putrescine or cadaverine) bound in "site B" to produce an imine intermediate that is reduced by NADH to release either homospermidine or aminobutylcadaverine. As the *E. coli* expression host used in this investigation naturally contains putrescine, cadaverine and spermidine, expression of the various HSS enzymes in *E. coli* leads to the production of diaminopropane or ammonia as a by-product when spermidine or putrescine was a substrate in "site A" and homospermidine or aminobutylcadaverine when putrescine or cadaverine was the co-substrate in "site B" (Figure 4.8).

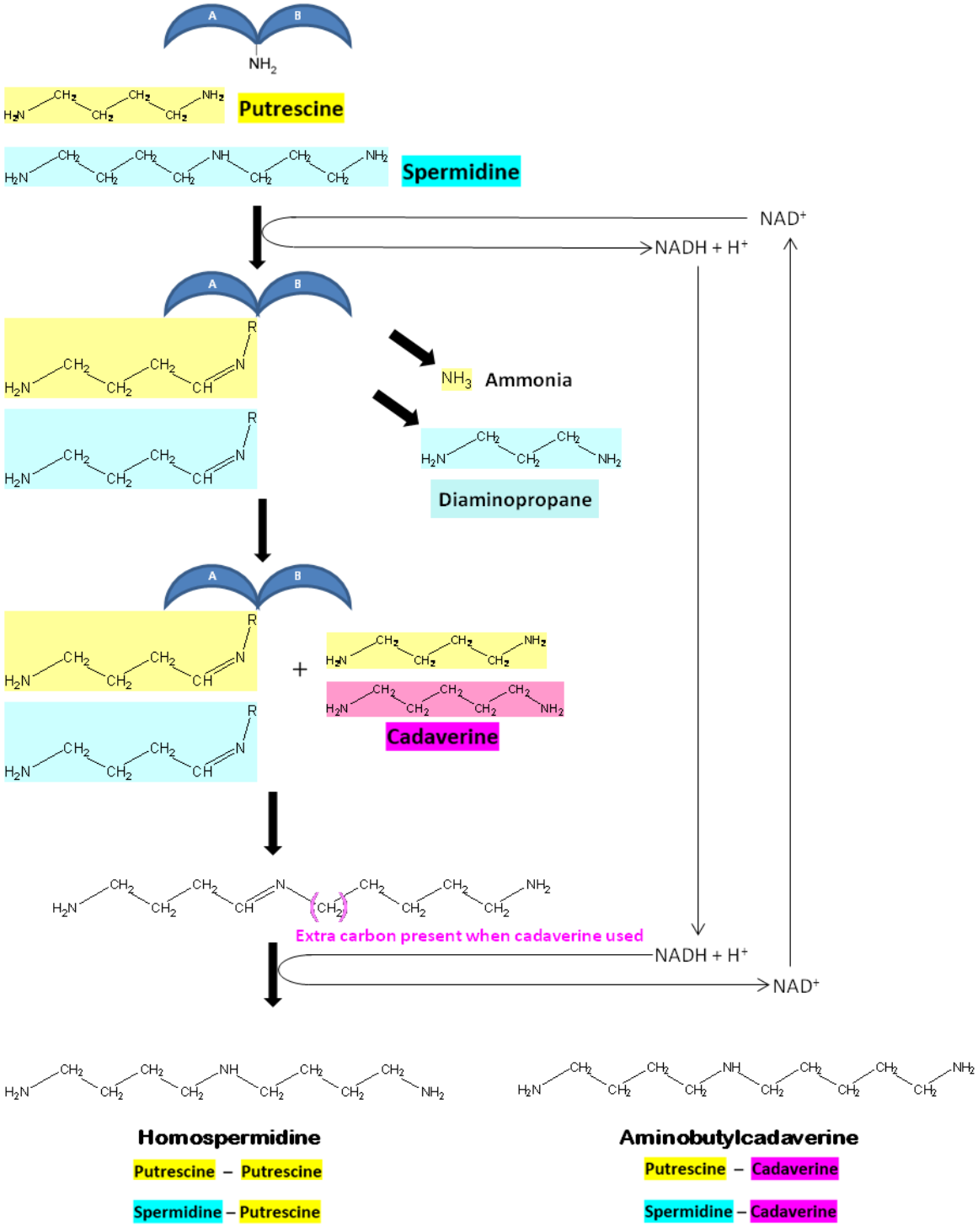


Figure 4.8 The production of homospermidine and aminobutylcadaverine from different substrates by bacterial HSS.
The possible composition of homospermidine and aminobutylcadaverine are shown at the bottom of the image. The diagram depicts the two binding sites of the HSS; A & B, and clearly shows the regeneration of NAD^+ during the reaction. [Adapted from 12].

Yammato *et al.* [14] showed the *in vitro* production of aminobutylcadaverine (but called it aminopentylputrescine) when they incubated the *Acinetobacter tartarogenes* HSS with cadaverine and putrescine. They believed it was created when a Schiff base formed between the aminobutylaldehyde and cadaverine. The production of the HSS by-products diaminopropane and aminobutylcadaverine was shown *in vitro* with purified cellular extracts of *R. viridis* [12], using putrescine (site A) and cadaverine (site B) as substrates.

The first natural occurrence of aminobutylcadaverine was reported by Fujihara *et al.* [15]. They concluded that the bacteria found in the root nodules of adzuki bean plant (*Vigna angularis*) are able to synthesise aminobutyl-cadaverine through the action of homospermidine synthase under a cadaverine-rich environment. Finally, Tholl *et al.* [13] eluded to the expression of *R. viridis* HSS in *E. coli* led to the production of aminobutylcadaverine (due to the presence of intracellular cadaverine) but these results were never published.

This is the first record of the production of aminobutylcadaverine during *in vivo* expression of several orthologues of HSS in *E. coli*. In addition, this is the first example of *in vivo* production of diaminopropane from spermidine by HSS. Previously, both of these products were predicted to have been formed using Böttcher's model [12], and were seen during his *in vitro* experiments, but the ability to make them *in vivo* has not been demonstrated previous to this investigation [17].

The ability of HSS to create new polyamine products using alternative substrates is interesting. The identification of naturally, or synthetically, occurring polyamines and polyamine conjugates are of constant interest, especially with respect to drug design and anti-tumoral properties.

4.5 Future Directions

The *B. japonicum* HSS enzyme has been characterised fully, but further investigations into the particular substrate affinities *in vitro* would be interesting experiments. The ability of *B. japonicum* HSS to use cadaverine as a substrate and form aminobutylcadaverine when heterogously expressed in *E. coli* is known (discussed in Section 4.4.3). However, no studies regarding the substrate affinity of the recombinantly purified enzyme have been performed. It would be novel to investigate if *B. japonicum* HSS has a higher affinity for cadaverine than putrescine, as cadaverine is not naturally found in *B. japonicum* (Figure 4.1).

In addition to further testing of the *B. japonicum* HSS, the other five *hss* orthologues could be purified and assayed *in vitro*. Assaying *in vitro* can provide information such as kinetic data, substrate specificity, substrate affinity, optimum temperature and pH. It would be interesting to determine if phylogenetically distinct *hss* orthologues possess different kinetic parameters or optimum conditions. As these genes are already translated with the N-terminally bound T7-tag in *E. coli*, the purification of the HSS enzymes could be done using the novogen T7 protein purification kit, which purifies target proteins using a rapid immunoaffinity purification method.

Chapter Four - References

1. Tabor CW, Tabor H: **Polyamines in microorganisms**. *Microbiological Reviews* 1985, **49**(1):81-99.
2. Pegg AE: **Recent Advances in the Biochemistry of Polyamines in Eukaryotes**. *Biochemical Journal* 1986, **234**(2):249-262.
3. Montanez R, Sanchez-Jimenez F, Aldana-Montes JF, Medina MA: **Polyamines: metabolism to systems biology and beyond**. *Amino Acids* 2007, **33**(2):283-289.
4. Wallace HM, Fraser AV, Hughes A: **A perspective of polyamine metabolism**. *Biochemistry Journal* 2003, **376**(1):1-14.
5. Wallace HM, Fraser AV: **Inhibitors of polyamine metabolism: review article**. *Amino Acids* 2004, **26**(4):353-365.
6. Hamana K, Matsuzaki S: **Occurrence of sym-homospermidine in the Japanese newt, *Cynops pyrrhogaster pyrrhogaster***. *FEBS Letters* 1979, **99**(2):325-328.
7. Hamana K, Matsuzaki S: **Widespread occurrence of norspermidine and norspermine in eukaryotic algae**. *Journal of Biochemistry* 1982, **91**(4):1321-1328.
8. Hamana K, Miyagawa K, Matsuzaki S: **Occurrence of Sym-Homospermidine as the Major Polyamine in Nitrogen-Fixing Cyanobacteria**. *Biochemical and Biophysical Research Communications* 1983, **112**(2):606-613.
9. Hamana K, Matsuzaki S, Sakakibara M: **Distribution of Sym-Homospermidine in Eubacteria, Cyanobacteria, Algae and Ferns**. *Fems Microbiology Letters* 1988, **50**(1):11-16.
10. Hamana K, Minamisawa K, Matsuzaki S: **Polyamines in Rhizobium, Bradyrhizobium, Azorhizobium and Argobacterium**. *FEMS Microbiology Letters* 1990, **71**(1-2):71-76.
11. Tait GH: **The formation of homospermidine by an enzyme from *Rhodopseudomonas viridis*** *Biochemical Society Transactions* 1979, **7**(1):199-201.
12. Böttcher F, Ober D, Hartmann T: **Biosynthesis of pyrrolizidine alkaloids: putrescine and spermidine are essential substrates of**

enzymatic homospermidine formation. *Canadian Journal of Chemistry* 1994, **72**(1):80-85.

13. Tholl D, Ober D, Martin W, Kellermann J, Hartmann T: **Purification, molecular cloning and expression in *Escherichia coli* of homospermidine synthase from *Rhodopseudomonas viridis***. *European Journal of Biochemistry* 1996, **240**(2):373-379.
14. Yamamoto S, Nagata S, Kusaba K: **Purification and Characterization of Homospermidine Synthase in *Acinetobacter tartarogenes Atcc-31105***. *Journal of Biochemistry* 1993, **114**(1):45-49.
15. Fujihara S, Abe H, Yoneyama T: **A new polyamine 4-aminobutylcadaverine. Occurrence and its biosynthesis in root nodules of adzuki bean plant *Vigna angularis***. *Journal of Biological Chemistry* 1995, **270**(17):9932-9938.
16. Michaelis L, Menten M: **Die Kinetik der Invertinwirkung**. *Biochemistry* 1913, **Z**(49):333-369.
17. Shaw FL, Elliott KA, Kinch LN, Fuell C, Phillips MA, Michael AJ: **Evolution and multifarious horizontal transfer of an alternative biosynthetic pathway for the alternative polyamine sym-homospermidine**. *Journal of Biological Chemistry* 2010, **285**(19):14711-14723.
18. Chin KJ, Liesack W, Janssen PH: ***Opitutus terrae* gen. nov., sp. nov., to accommodate novel strains of the division 'Verrucomicrobia' isolated from rice paddy soil**. *International Journal of Systematic and Evolutionary Microbiology* 2001, **51**(6):1965-1968.
19. O'Kelly C, Silberman J, Amaral Zettler L, Nerad TA, Sogin ML: ***Monopylocystis visvesvarai* n. gen., n. sp. and *Sawyeria marylandensis* n. gen., n. sp.: Two New Amitochondrial Heterolobosean Amoebae from Anoxic Environments**. *Protist* 2003, **154**(2):281-290.
20. Aury JM, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, Segurens B, Daubin V, Anthouard V, Aiach N *et al*: **Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia***. *Nature* 2006, **444**(7116):171-178.
21. Lee J, Michael AJ, Martynowski D, Goldsmith EJ, Phillips MA: **Phylogenetic diversity and the structural basis of substrate**

specificity in the beta/alpha-barrel fold basic amino acid decarboxylases. *Journal of Biological Chemistry* 2007, **282**(37):27115-27125.

22. Williams KP, Sobral BW, Dickerman AW: **A robust species tree for the alphaproteobacteria**. *Journal of Bacteriology* 2007, **189**(13):4578-4586.
23. Gupta RS, Mok A: **Phylogenomics and signature proteins for the alpha proteobacteria and its main groups**. *BMC Microbiology* 2007, **7**:106.
24. Scherer P, Kneifel H: **Distribution of polyamines in methanogenic bacteria**. *Journal of Bacteriology* 1983, **154**(3):1315-1322.
25. Deppenmeier U, Johann A, Hartsch T, Merkl R, Schmitz RA, Martinez-Arias R, Henne A, Wiezer A, Baumer S, Jacobi C *et al*: **The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea**. *Journal of Molecular Microbiology and Biotechnology* 2002, **4**(4):453-461.
26. Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S, Wu D, Eisen JA, Hoffman JM, Remington K *et al*: **The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific**. *PLoS Biology* 2007, **5**(3):e77.
27. Hamana K, Takeuchi M: **Polyamine profiles as chemotaxonomic markers within alpha, beta, gamma, delta, and epsilon subclasses of class Proteobacteria: distribution of 2-hydroxyputrescine and homospermidine.** *Microbial Culture Collection* 1998, **14**:1-14.

Chapter Five

Physiological Function of Polyamines

TABLE OF TABLES

TABLE OF FIGURES

Chapter Five - Physiological Function of Polyamines

5.1 Background

The exact role and nature of polyamines differs between species, but we know that they are essential for normal growth and development in organisms [1]. Polyamines are found in nearly all organisms on the planet, but interestingly, an identical polyamine profile is not universally necessary, not even in closely related species. For example, in the family Intrasporangiaceae, the genus *Janibacter* contains putrescine and cadaverine, whereas *Intrasporangium* and *Terracoccus* (also from Intrasporangiaceae) both contain putrescine and spermidine, and further to this, *Terrabacter* contains only putrescine [2].

The existence of organisms that use only asymmetrical spermidine or symmetrical homospermidine as their sole triamine has profound implications for the molecular mechanisms of action of polyamines. It is known that polyamines are able to bind to DNA, and have a wide range of roles within the cell, so it is particularly remarkable that these polyamines come in a variety of shapes and sizes. Even with different charge states and structures, the same roles are fulfilled.

For this study, *Rhizobium leguminosarum* and *Campylobacter jejuni* were chosen as representative organisms from the α- and ε-proteobacteria respectively to investigate the specificity of the polyamine profile. *R. leguminosarum* and *C. jejuni* are species with established optimal growth conditions. *C. jejuni* is responsible for causing of human gastroenteritis and its pathogenesis and survival are common topics of investigation. No studies on *C. jejuni* polyamines exist at the time of writing, except one paper describing the crystal structure of the *C. jejuni* biosynthetic arginine decarboxylase (ADC) [3]. *R. leguminosarum* is a root-nodule forming bacteria and has been studied extensively [4, 5]. *R. leguminosarum* has previously been examined for polyamine content. Some fast growing strains of Rhizobia have been reported as containing putrescine, spermidine and aminobutylhomospermidine [6]. The complete genome sequences are available (Genbank Accession Numbers: *C. jejuni* –CP000814 [7], *R. leguminosarum* - AM236080 [8]).

5.2 Objectives

- To investigate the effect of a polyamine pathway inhibitor (DFMO) on *R. leguminosarum.*

- To understand if the exact structure of the triamine is necessary for normal *R. leguminosarum* growth using exogenous polyamines (native and non-native) and DFMO, the ODC inhibitor.

- To identify the polyamines naturally present in *C. jejuni* and predict the biosynthetic enzymes involved.

- To investigate if the normal triamine in *C. jejuni* can be replaced by an alternative non-native triamine using exogenous polyamines and genetic manipulation.

5.3 Results

5.3.1 Homospermidine is required for normal growth in the α-Proteobacterium *R. leguminosarum*

R. leguminosarum has a simple polyamine pathway which includes the synthesis of putrescine from ornithine by an eukaryotic-like alanine racemase fold ornithine decarboxylase (ODC), followed by homospermidine production from putrescine by homospermidine synthase (HSS) (Figure 3.5). *R. leguminosarum* was grown in the presence of the ODC inhibitor, D,L-α-difluoromethylornithine (DFMO). In principle, inhibition of the ODC enzyme in *R. leguminosarum* should decrease putrescine content and consequently, homospermidine levels.

DFMO was added to *R. leguminosarum* defined minimal growth medium (AMS). The cells were harvested and resuspended in an appropriate buffer and the ODC activity of the live samples was tested using a radioactive assay. The *R. leguminosarum* cells grown aerobically in the presence of 1.0 mM DFMO were found to have a specific ODC activity of 1.82 ± 0.03 nmol CO_2 h^{-1} mg^{-1} of protein when compared with 4.29 ± 0.01 nmol CO_2 h^{-1} mg^{-1} of protein for the control cultures grown without DFMO. These results show that 1.0 mM DFMO exogenously supplied causes a more than two-fold decrease in *R. leguminosarum* ODC activity (Figure 5.1).

Growing *R. leguminosarum* cells in increasing concentrations of DFMO resulted in correspondingly decreased growth rates (Figure 5.2). Colony forming unit counts were not performed on these cells, however the cell morphology was inspected under the microscope to determine that the cell shape remained unaltered after growth in DFMO. Therefore the change in optical density (OD) was directly proportional to the amount of cells, not an artefact of differential cell morphology. When *R. leguminosarum* were grown in 5 mM DFMO, growth rate appeared to decrease by up to four-fold. With concentrations of DFMO of more than 5 mM, the growth rate appeared to remain similar, so 5 mM DFMO was used for subsequent experiments.

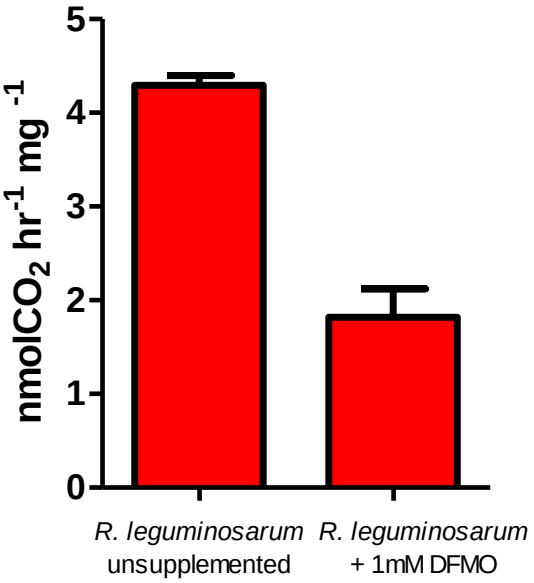


Figure 5.1 The *R. leguminosarum* ODC activity.
Cells are grown in 1 mM DFMO and control AMS media. Standard deviation shown based on three biological repeats.

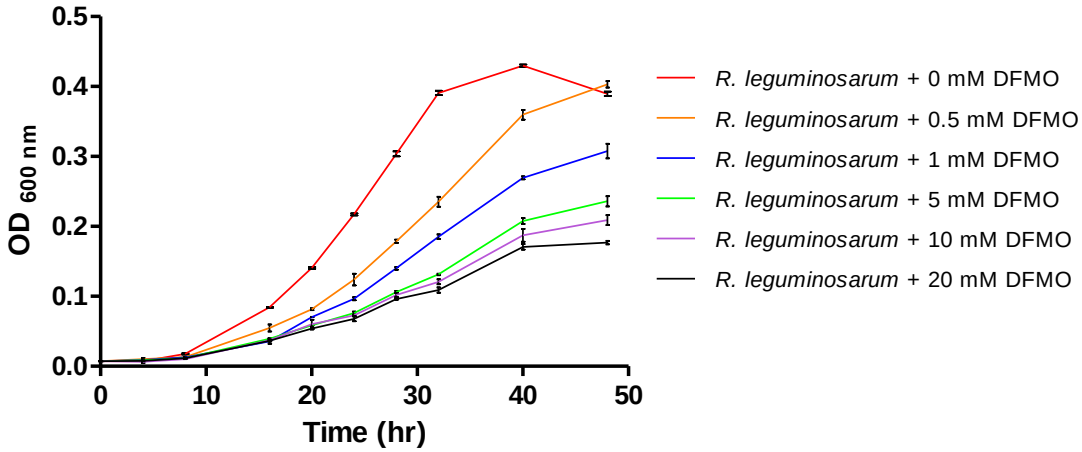


Figure 5.2 The effect of DFMO on *R. leguminosarum* growth.
Measured by OD_{600nm}, with increasing DFMO concentrations (see legend) added to the growth medium (AMS). Standard deviation shown based on three biological repeats.

The polyamine content of *R. leguminosarum* cells grown aerobically with or without 5 mM DFMO is shown in Figure 5.3. After 24 hours growth, there was a dramatic decrease in putrescine concentration, from 1.6 to 0.1 (arbitrary units based on the normalised HPLC peak area), and a 20-fold decrease in homospermidine concentration (from 34.8 to 1.7) in the presence of the inhibitor, DFMO. There was a more severe decrease in putrescine (from 14.3 to 0.5) and homospermidine (from 90.8 to 5.1) in the presence of the inhibitor after 48 hours of growth.

There was a three-fold decrease in diaminopropane content (from 15.7 to 5.6, after 48 hours) seen in between those cells grown in the presence or absence of 5 mM DFMO. This change may have been due to changes in growth rate, as diaminopropane concentration should not be directly affected by DFMO inhibition of ODC activity. Currently, the occurrence of intracellular diaminopropane is unexplained, but it is understood that the levels of intracellular polyamines change with respect to growth rate [9].

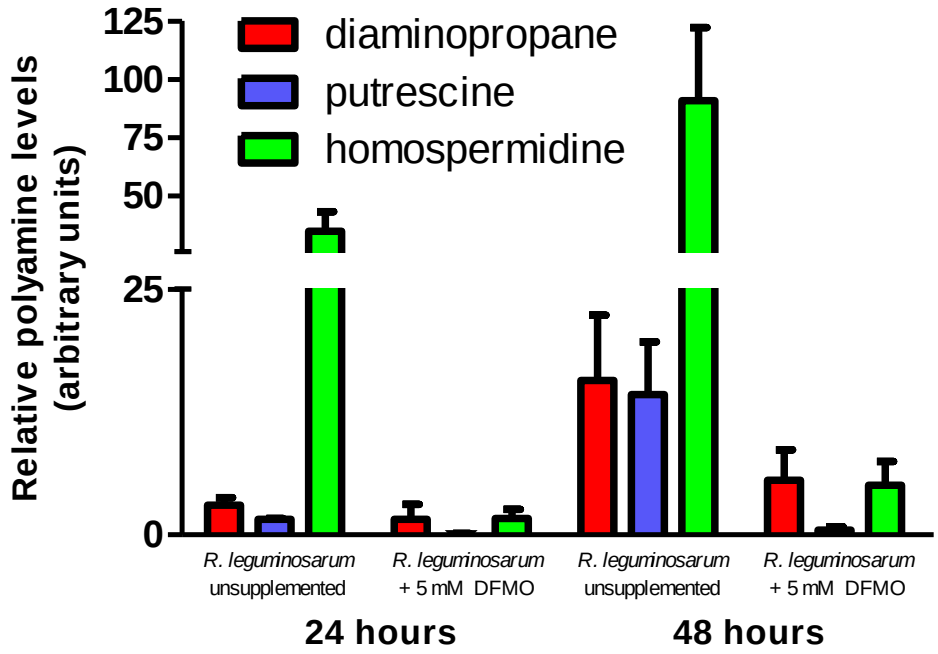


Figure 5.3 Quantified intracellular polyamine content of *R. leguminosarum* grown in the absence and presence of DFMO (5 mM). Sampled at two time points – 24 & 48 hours. The data was calculated by the HPLC peak area (mAU) corrected for the internal standard (Diaminononane) and cell pellet weight (mg). The standard error is shown based on 3 biological replicates.

Exogenous polyamines restoring *R. leguminosarum* growth: It was possible that adding a potent inhibitor (DFMO) in such high concentrations to the growth medium could have had adverse effects on the cells metabolism and could have possibly altered functions such as DNA replication and cell division. To verify that the decrease in growth rate caused by DFMO was due to polyamine depletion, *R. leguminosarum* cells were grown aerobically in 5 mM DFMO-containing medium supplemented with 1 mM putrescine. As seen in Figure 5.4, putrescine restores growth of DFMO-treated cells almost completely, confirming that the effect of DFMO on cell growth is due to polyamine depletion. It can be hypothesised that if putrescine is present, the rest of the pathway will function as usual and homospermidine will be produced. One this note, HPLC of cell extracts confirmed the uptake of putrescine by *R. leguminosarum*, as it was seen intracellularly when compared

to the DFMO-treated cell extracts. In addition, the restoration of intracellular homospermidine levels was observed (Figure 5.5).

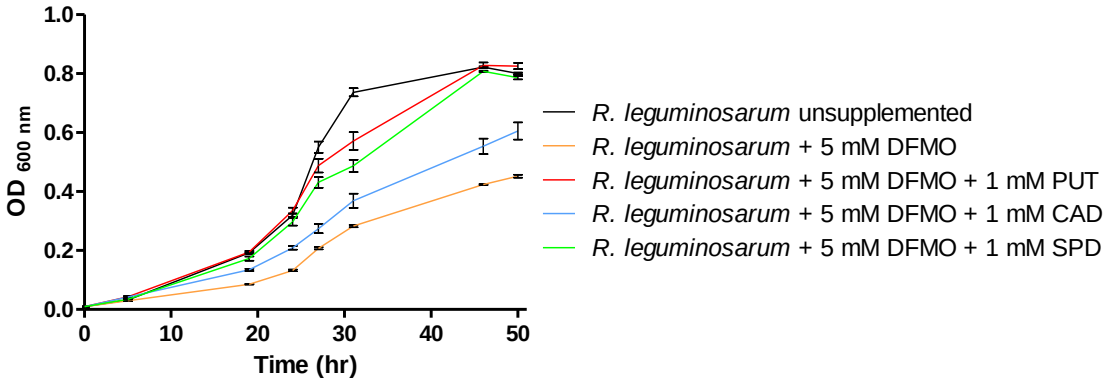


Figure 5.4 Growth of *R. leguminosarum* in minimal AMS medium containing DFMO and exogenous polyamines.
(see legend), PUT, putrescine; CAD, cadaverine; SPD, spermidine. Standard deviation shown based on three biological repeats.

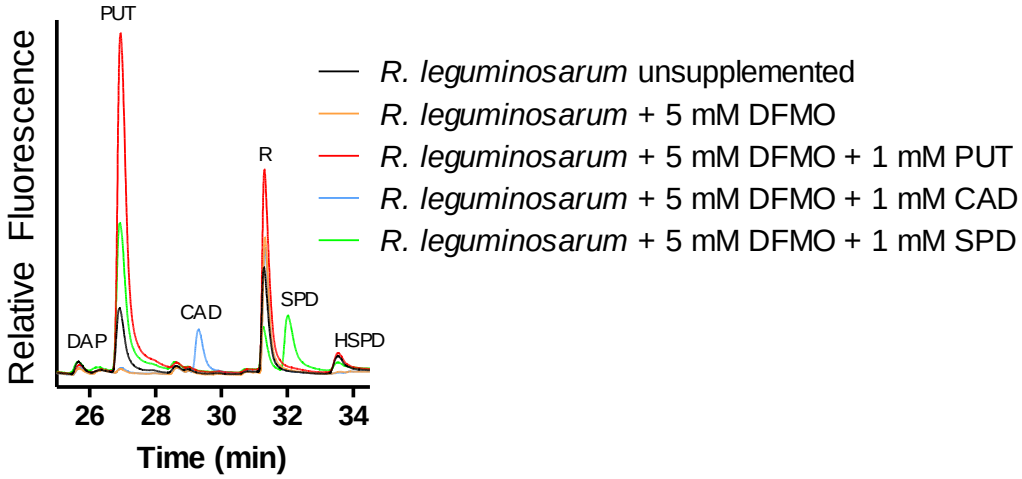


Figure 5.5 The intracellular polyamine profile, detected by HPLC, of *R. leguminosarum* wild-type cells grown in minimal AMS medium containing DFMO and exogenous polyamines
(see legend), DAP, diaminopropane; PUT, putrescine; CAD, cadaverine; SPD, spermidine; R, label. HPLC internal standard and polyamine standards not shown.

Another diamine, cadaverine (1 mM, Figure 3.2), was much less effective than putrescine in restoring growth of *R. leguminosarum* whilst grown in 5 mM DFMO (Figure 5.4), although cadaverine was found to be taken up and accumulated in the cells (identified by HPLC, Figure 5.5). This suggests that cadaverine is not able to replace putrescine as the major diamine in *R. leguminosarum*. Bacterial HSS is able to use cadaverine as a substrate, along with putrescine, to produce aminobutylcadaverine [10]. But in this case, HPLC analysis of the intracellular polyamines did not show the accumulation of aminobutylcadavine. The lack of aminobutylcadaverine was probably because putrescine was so effectively depleted by DFMO (Figure 5.3), and therefore could not combine with the cadaverine. Cadaverine alone cannot be used as a substrate for HSS [11] and this was confirmed by HPLC of the intracellular polyamines, as homospermidine, or another possible HSS product, was not produced (Figure 5.5).

Interestingly, spermidine was able to restore growth of the DFMO-treated cells almost as well as putrescine (Figure 5.4). The restored growth was accompanied by the intracellular identification of homospermidine and putrescine as shown by HPLC analysis (Figure 5.5). The presence of putrescine could be due to incomplete inhibition of the ODC by DFMO, but as this was not seen in the negative control (DFMO only) it is unlikely. It is possible that the putrescine is being formed from the breakdown of additional exogenous spermidine. Homospermidine is also present intracellularly (as shown by HPLC analysis, Figure 5.5), and it must have been formed from the spermidine or putrescine present, by the HSS enzyme. It has been shown *in vitro* that HSS is able to produce homospermidine, putrescine and diaminopropane from spermidine alone [11].

Exogenous triamines restoring *R. leguminosarum* growth: The ability of three different triamines: norspermidine, spermidine and homospermidine to restore growth of DFMO-treated *R. leguminosarum* cells was compared. Growth was inhibited by 5 mM DFMO, and each triamine was added to the minimal AMS growth media. All three of the triamines restored growth to a similar level (Figure 5.6), and each triamine accumulated in cells to a similar level (confirmed by HPLC, Figure 5.7).

As before, exogenous spermidine was taken up and converted by the HSS into homospermidine, putrescine and diaminopropane (Figure 5.7). The exogenous homospermidine was taken up by the *R. leguminosarum* (Figure 5.7). The exogenous norspermidine was taken up, but not converted to homospermidine (Figure 5.7). This result was interesting as it confirmed that *R. leguminosarum* does not require its natural, specific triamine, homospermidine, to function. It appeared to grow just as well with norspermidine, which is symmetrical, like homospermidine, but two carbons shorter in length, or spermidine, which is an asymmetrical triamine (Figure 3.2). This is interesting because *R. leguminosarum* has evolved to produce only homospermidine, so it raises the question: how and why does the cell function apparently easily with norspermidine or spermidine instead? In nature, the rhizobia may be able to utilise a selection of triamines found naturally in the environment, but when there are none available, the species are most likely stimulated into synthesising their own. Taking-up polyamines from the surrounding environment would most likely be a lower energy process for the micro-organisms, than synthesising their own.

To directly compare the effects of exogenous spermidine, norspermidine and homospermidine on growth in *R. leguminosarum*, it will be necessary to have a gene deletion of *hss* in this species so that spermidine is not metabolised to homospermidine by HSS. Then, true comparisons of whether the size and symmetry of the triamine in *R. leguminosarum* could be made.

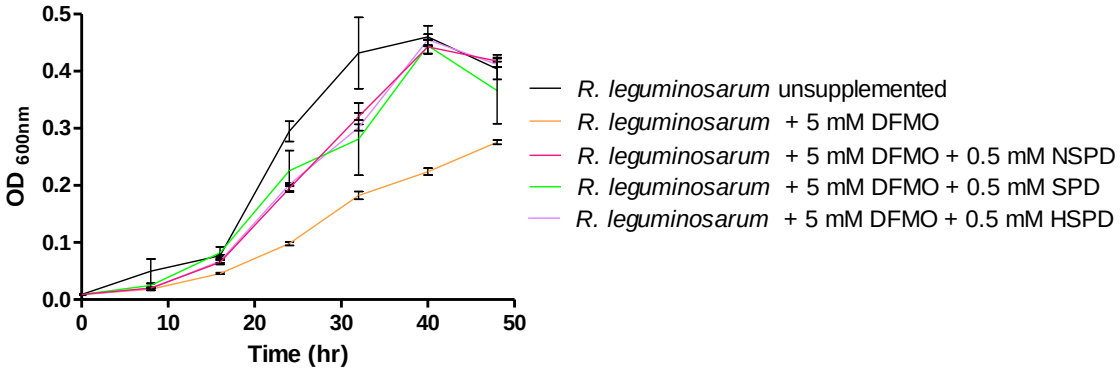


Figure 5.6 Growth of *R. leguminosarum* detected by OD_{600nm}, in a minimal AMS medium with DFMO and exogenous triamines.
(see legend) HSPD, homospermidine; NSPD, norspermidine; SPD, spermidine. Standard deviation shown based on three biological repeats.

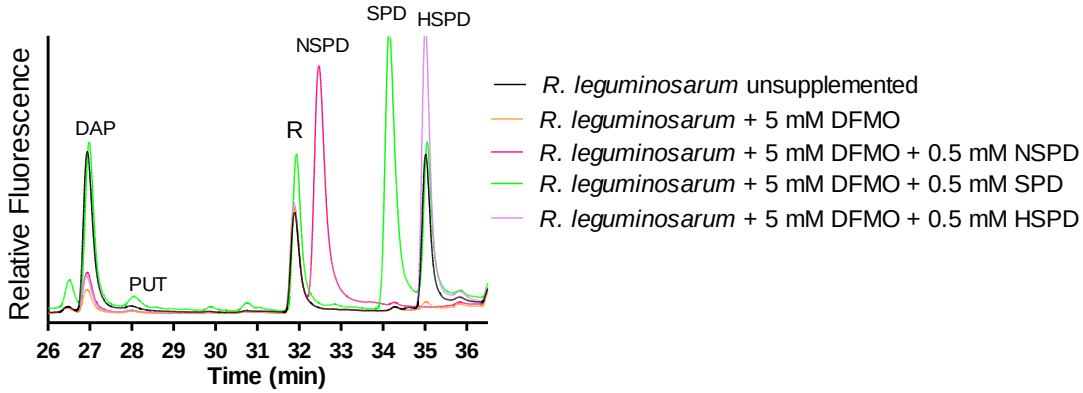


Figure 5.7 Intracellular polyamine profile, detected by HPLC, of *R. leguminosarum* cells grown with DFMO and exogenous polyamines.
(see legend) DAP, diaminopropane; PUT, putrescine; R, label; NSPD, norspermidine; SPD, spermidine; HSPD, homospermidine.

NOTE: The elution time of spermidine changed during this investigation because a new column was purchased for HPLC analysis. The new column, although the same as the previous, possessed a slightly different retention property for spermidine. The peaks were confirmed by the use of polyamine standards in each run.

5.3.2 Characterisation of the KO-100 *R. leguminosarum*

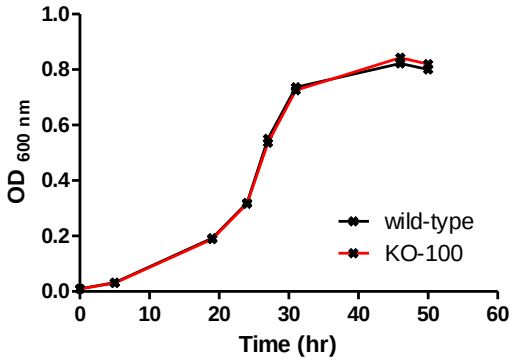
A mutant with a transposon (Tn5) insertion 100 bp upstream of the translational start codon of *hss* in *R. leguminosarum* has been identified and termed 'KO-100'. The Tn5 mutant library was constructed at the John Innes Centre, Norwich, by Adam Wilkinson and transposon insertion mutants were located in the Tn5 library of *R. leguminosarum* strain 3841 using PCR primers complementary to the end of transposon and the gene of interest as described by Fox *et al.* [12]. The transposon is most likely located in the promoter region of the *hss*, and the regulation of the gene has probably been affected by the transposon insertion.

This insertion mutant has been phenotypically analysed for changes in growth and polyamine metabolism. Investigation showed that the KO-100 and wild-type *R. leguminosarum* were able to grow at similar rates; however, the KO-100 contained nearly 3-fold less homospermidine than the wild-type (quantified and identified by HPLC, Figure 5.8). This suggests that under normal conditions, *R. leguminosarum* contains more homospermidine than appears necessary for a phenotypically normal cell. The role of homospermidine in *R. leguminosarum* is currently unknown, although suggestions that it has a role in nitrogen fixation have been made based on homospermidine taxonomic distribution [13]. It is possible that the large reserve of homospermidine could be stored in *R. leguminosarum* until required.

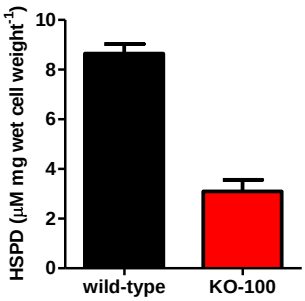
In addition to the decrease in homospermidine, there also appeared to be an accumulation of putrescine, the substrate used to synthesise homospermidine in *R. leguminosarum* (Figures 3.5 and 5.8c). This suggests that the HSS activity is decreased, and so less homospermidine is synthesised from putrescine.

The KO-100 was inoculated onto the pea plant, but no distinct phenotype regarding nodule formation was observed (Figure 5.9). Microscopy of the KO-100 also proved inconclusive. The KO-100 did not appear to have any distinct physical phenotype (Figure 5.10). This suggests that this Tn5 insertion (100 bp upstream of the translational start codon of hss) in *R. leguminosarum* is not detrimental to the species.

a



b



c

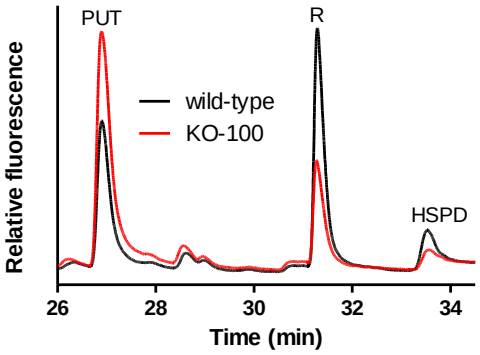


Figure 5.8 Characterisation of KO-100.
(a) The growth of KO-100 compared to the wild-type *R. leguminosarum* based on three biological replicates (b) The amounts of intracellular homospermidine quantified by HPLC the standard error based on the three biological repeats (c) HPLC trace of internal polyamines in KO-100 and wild-type *R. leguminosarum:* PUT, putrescine; R, label; HSPD, homospermidine.

Figure 5.9 Nodule formation on pea plants.
Inoculation with wild-type *R. leguminosarum* (left) and KO-100 (right). Nodulation was similar on both plants.

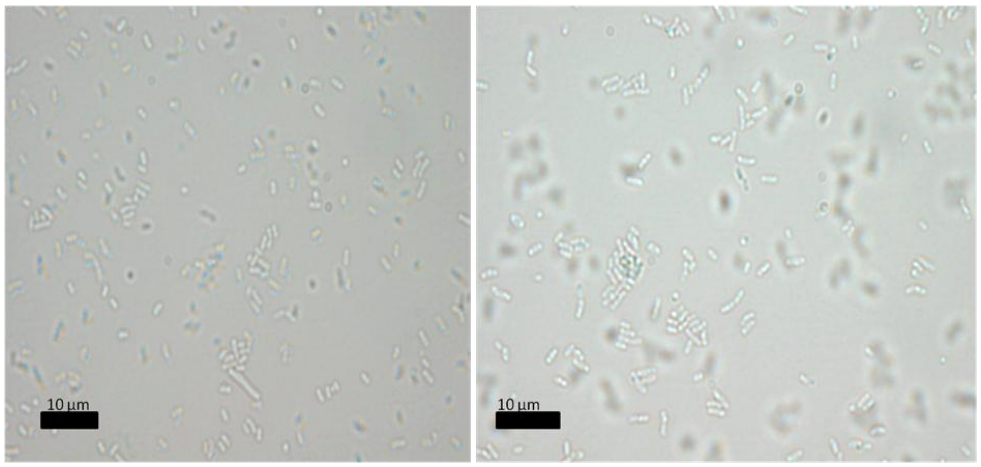


Figure 5.10 Light microscopy of *R. leguminosarum.*
Wild-type *R. leguminosarum* (left) and KO-100 (right). At this magnification, no differences were seen between the samples.

Extensive screening was performed to try and locate a *R. leguminosarum* mutant with the Tn5 transposon inserted inside the *hss*, but unfortunately, one was never identified. The inability to identify a *hss* mutant suggested that a mutation in this gene may be lethal for *R. leguminosarum*. If the gene is essential for survival, any alteration within the transcript of the gene could result in non-viable cells. This would make the mutation impossible to identify using this screening method. However, a study which used *Rhizobium etli* in a search for novel genetic determinants of swarming used 700 miniTn5 mutants, one of which was a homospermidine synthase knock-out. The mutant was viable and able to swim like the wild-type [14]. This argues that a complete *hss* inactivation is possible in the Rhizobium species. Currently, *R. leguminosarum* and *R. etli* are both annotated as containing only one *hss* gene, and there is no known or reported gene duplication of *hss* in this species. The experiments were performed using medium with no additional polyamine supplements, however the medium used most likely contained significant amounts of polyamines.

5.3.3 *C. jejuni* requires a triamine for normal growth, however the exact structure of the triamine is not essential

There are only a minority of studies involving polyamines in the ε-Proteobacteria. *Helicobacter pylori* has been studied for polyamine content [15]. This study showed the major triamine in *H. pylori* to be spermidine. Hamana *et al.* [16] described the polyamine content of Campylobacter as scarcely containing any polyamines, and there are no reports on the Campylobacter polyamine pathway, so *C. jejuni* was chosen as the model micro-organism from the ε-Proteobacteria for this investigation. The crystal structure of arginine decarboxylase (ADC) has been determined in *C. jejuni* [3] and has been compared to that of *E. coli*.

The possible polyamine pathway was identified based on the sequence homology of certain *C. jejuni* genes to other annotated polyamine biosynthetic genes (refer to Figure 3.4). This suggested that *C. jejuni* did not contain the genes to synthesise spermidine in the same way as the closely related *H. pylori*. Instead of a spermidine synthase (like *H. pylori*), *C. jejuni* is predicted to

contain a carboxyspermidine decarboxylase (CANSDC). The possibility of an unusual polyamine biosynthetic pathway, coupled with the ease of genetic manipulation of *C. jejuni* in the laboratory made it a particularly attractive candidate for this investigation.

The polyamines were extracted from *C. jejuni*, analysed by HPLC and quantified based on peak area and appropriate polyamine standard curves. Polyamine standards were run alongside the *C. jejuni* sample, and the polyamines were identified based on their elution times. This confirmed the intracellular presence of spermidine, but not putrescine (Figure 5.11 and Table 5.1). It is possible that as an intermediate, putrescine is formed and metabolised instantly (Figure 3.4); this would explain the lack of putrescine. The relatively low amount of spermidine observed at stationary growth phase (Table 5.1) could explain why previous reports detected no polyamines in *C. jejuni* [16], especially if the detection method used was not as sensitive as described here and the cells tested where at stationary growth phase.

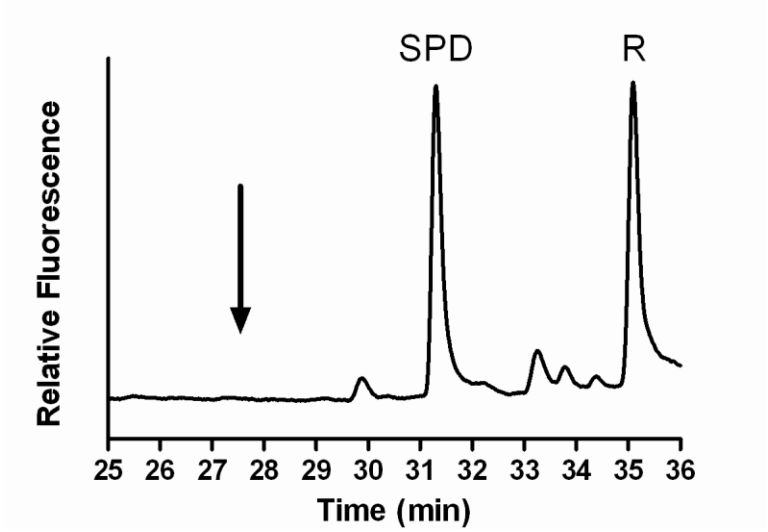


Figure 5.11 A HPLC Chromatograph of *C. jejuni.*
Analysis by Colin Hanfrey
The polyamine content of *C. jejuni* wild-type strain 81116. R, label; SPD, spermidine; the arrow represents the area that putrescine is expected to be seen.

Growth phase	Putrescine	Norspermidine	Spermidine
	nmoles.g^{-1} cell fresh weight		
Log (OD_{600nm} 0.2 to 0.3)	n.d.	n.d.	865.8 ± 192.6
Stationary (OD_{600nm} 0.9 to 1.0)	n.d.	n.d.	296.5 ± 112.0

Table 5.1 Cellular polyamine content of *C. jejuni* wild-type strain 81116.
Analysis by Colin Hanfrey
Values are means of 3 replicates ± standard error. n.d., not detectable.

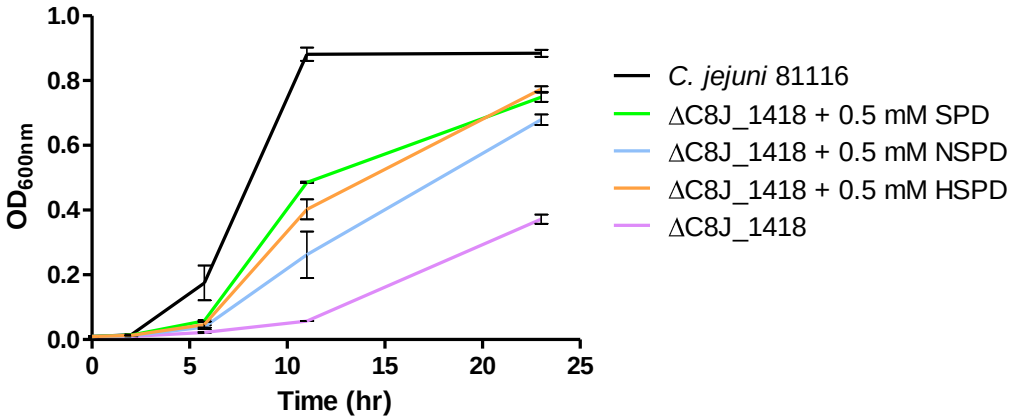
A mutant was made in the putative carboxyspermidine decarboxylase (CANSDC, the gene reference number C8J_1418) shown in Figure 3.4. The CANSDC is the final enzyme in the predicted polyamine pathway in *C. jejuni*. The polyamine carboxyspermidine is decarboxylated by CANSDC to make the triamine, spermidine. The insertional activation mutant was created by replacing the CANSDC gene with a kanamycin resistance cassette.

CANSDC mutant phenotype and exogenous polyamines: The CANSDC mutant (ΔC8J_1418) had a severely decreased growth phenotype. This supports the idea that polyamines are crucial for normal cellular growth. The defected-growth phenotype could be partially restored by the exogenous addition of 0.5 mM spermidine (the native triamine), 0.5 mM norspermidine or 0.5 mM homospermidine (both triamines not found naturally in *C. jejuni*) (Figure 5.12a). HPLC of the intracellular polyamines of *C. jejuni* showed that norspermidine, homospermidine and spermidine were all taken up by the micro-organism (Figure 5.12b). This means that *C. jejuni* must have transporters capable of importing these triamines. These data suggest that a triamine improves *C. jejuni* growth, but the exact shape of the triamine is not crucial. The function, or some of the functions, which spermidine performs in *C. jejuni* can be replaced by either homospermidine or norspermidine.

Complementation with mutated gene: The CANSDC mutant (ΔC8J_1418) was genetically complemented with the C8J_1418 gene under a native *C. jejuni* promoter; *fdx*A (compΔC8J_1418). The *fdx*A promoter was chosen for manipulated gene expression in *C. jejuni* because it is has been characterised previously, and is known to be highly induced under certain conditions (personal communication Duncan Gaskin, [17]). This CANSDC comple-

mentation re-established the defective growth phenotype, and the presence of intracellular spermidine was restored (Figure 5.13).

a



b

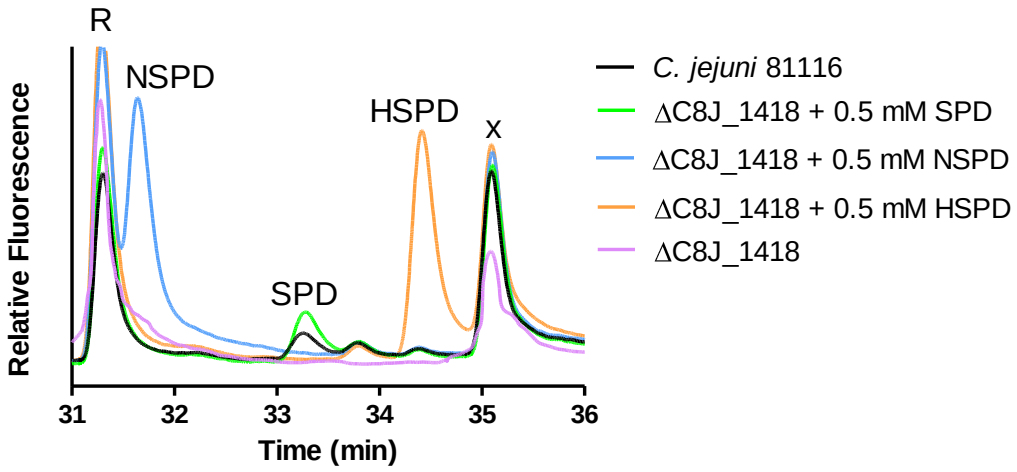
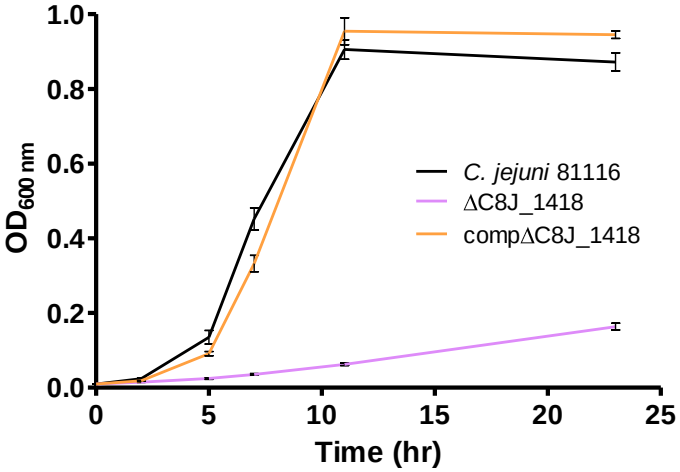


Figure 5.12 *C. jejuni* growth and polyamine content (polyamine complement)
Analysis by Colin Hanfrey
(a) The growth of *C. jejuni* wild-type (81116) and ΔC8J_1418 with the addition of exogenous polyamines (see legend) measured by OD_{600nm}. Standard deviation shown based on three biological repeats.
(b) The HPLC chromatograph shows the intracellular polyamines: X, internal standard; R, label; NSPD, norspermidine; SPD, spermidine; HSPD, homospermidine.

Complementation with alternative polyamine synthetic gene: Genetic complementation of the CANSDC mutant using the *Paramecium tetraurelia* homospermidine synthase (PARA-HSS), under a native Campylobacter promoter (*met*K), in the hope to restore growth using a non-native *in vivo* synthesised triamine was unsuccessful (ΔC8J_1418::PARA-HSS). The *met*K promoter was chosen for manipulated gene expression in *C. jejuni* because it is has been characterised previously, and is known induce genes to a relatively low level (personal communication Duncan Gaskin, [18]). Expression at a low level was chosen because the ability of *C. jejnui* to synthesise the PARA-HSS was unknown, and if unusual amino acids or resources were required, the hetrologous expression of the gene could render the host organism unviable. Although the genetic complementation worked (verified by PCR), and the gene was expressed (verified by RT-PCR, Figure 5.14a), the mutant was not viable in minimal medium. In rich medium, the cells were able to grow, but polyamine analysis by HPLC showed there was no intracellular homospermidine present (Figure 5.14).

The PARA-HSS gene was able to produce homospermidine *in vivo* when expressed in *E. coli* under the pET21 system, but was not able to synthesise homospermidine when expressed in *C. jejuni*. It is possible that the correct substrate, co-factors and the internal environment of *C. jejuni* were not suitable or available for the PARA-HSS to form homospermidine. The PARA-HSS was not assayed *in vitro*, so optimum conditions of PARA-HSS are still to be established.

a



b

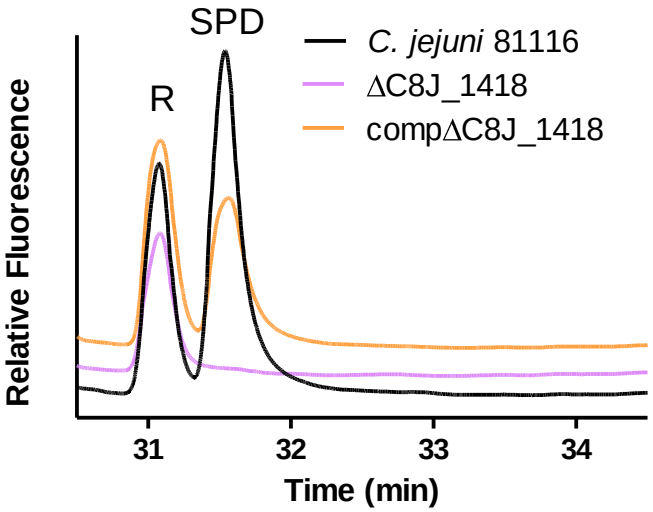
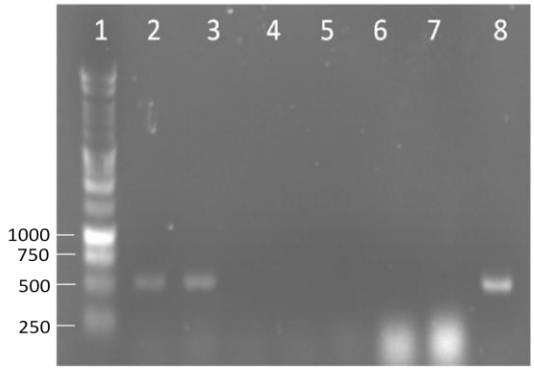


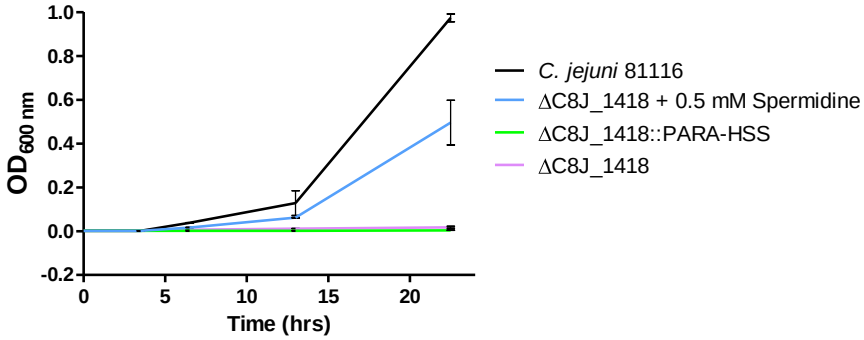
Figure 5.13 *C. jejuni* growth and polyamine content (genetic complement)
Analysis by Colin Hanfrey
(a) The growth of *C. jejuni* wild-type, ΔC8J_1418 and complemented mutant (compΔC8J_1418) measured by OD_{600nm}.
(b) HPLC chromatograph of the intracellular polyamines of *C. jejuni* wild-type, ΔC8J_1418 and complemented mutant (compΔC8J_1418); R, label; SPD, spermidine.

NOTE: The elution time of spermidine changed during this investigation because a new column was purchased for HPLC analysis. The new column, although the same as the previous, possessed a slightly different retention property for spermidine. The peaks were confirmed by the use of polyamine standards in each run.

a



b



c

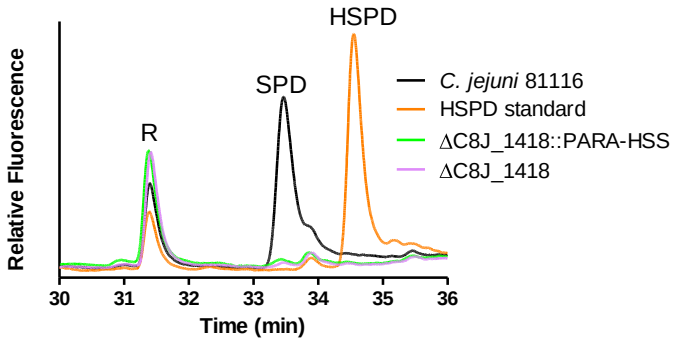


Figure 5.14 PARA-HSS expression in *C. jejuni*

(a) RT-PCR showing PARA-HSS expression. Lanes: 1 = ladder, Promega 1Kb, 2 & 3 = RT-PCR on ΔC8J_1418::PARA-HSS, 4 to 7 = negative controls, 8 = positive control.

(b) The growth of *C. jejuni*, ΔC8J_1418, ΔC8J_1418::PARA-HSS and the mutant chemically complemented with spermidine (ΔC8J_1418 + 0.5 mM Spermidine) measured by $OD_{600nm.}$ (Starting OD 0.01). Standard deviation shown based on three biological repeats.

(c) HPLC chromatograph of the intracellular polyamines of *C. jejuni,* ΔC8J_1418, and ΔC8J_1418::PARA-HSS (see legend). The homospermidine standard (HSPD - orange) is shown for comparison; SPD, spermidine; R, label.

5.4 Discussion

Using the bacterial model organisms *R. leguminosarum* and *C. jejuni*, several conclusions regarding the physiological function of polyamines have been drawn. This investigation is far from complete, but this study has revealed several areas for further exploration and development.

R. leguminosarum

The study has shown that 0.5 mM DFMO added to AMS minimal media can decrease *R. leguminosarum* growth. Increasing the concentration of DFMO (up to 20 mM) resulted in a correspondingly decreased growth rate. This is the first report to show that DFMO can inhibit bacterial growth in a polyamine dependent manner, and the finding has interesting implications for the potential effect of DFMO on bacterial pathogens that contain only the eukaryotic-like alanine racemase (AR) fold ODC-encoding gene for putrescine biosynthesis like *R. leguminosarum*. Such pathogens include the oral pathogen *Treponema denticola*; the agent of Q fever, *Coxiella burnetii*; *Brucella* species, which cause brucellosis, and *Bartonella* species, which cause diseases such as cat scratch fever. In principle, growth of each of these pathogens should be susceptible to inhibition by DFMO, and there may be antibacterial potential for this anticancer [19], antitrypanosomal [20] drug which is clinically approved for use in humans. It should be noted that DFMO is not active against the aspartate aminotransferase (ATT) fold ODC found in some bacterial species, like *E. coli* [21]. DFMO causes a growth defect in *H. pylori*, but this is a polyamine independent effect [22].

In addition to the decreased *R. leguminosarum* growth rate due to DFMO treatment, the intracellular polyamine profile was analysed by HPLC, and there was no intracellular putrescine or homospermidine in the DFMO-treated cells (Figure 5.5), both polyamines which are naturally found in *R. leguminosarum*. This confirmed that DFMO was inhibiting ODC, so putrescine, and consequently homospermidine, were not synthesised (Figure 3.5). An assay of the ODC activity in living *R. leguminosarum* cells after growth in DFMO (1 mM) showed there to be substantially less ODC activity in DFMO-treated cells when compared to the wild-type *R. leguminosarum* (Figure 5.1), which further confirmed that the DFMO is inhibiting the ODC.

The growth inhibition of *R. leguminosarum* caused by DFMO, can be reversed by the exogenous application of polyamines. In future phenotypic and genotypic investigations of *R. leguminosarum*, in addition to testing cells grown in the absence or presence of DFMO-supplemented media, those grown with exogenous DFMO and a polyamine supplement would be an excellent investigative sample to study because the polyamine profile appears to be easily manipulated. Even at 20 mM DFMO, a high physiological concentration, *R. leguminosarum* cultures were still able to grow. Interestingly, supplementation with the native polyamines (putrescine or homospermidine) almost fully restores growth. Also, non-native polyamines, such cadaverine, spermidine and norspermidine, can partially restore growth.

The ability of three triamines: spermidine, norspermidine and homospermidine to restore growth of DFMO-treated *R. leguminosarum* cells was demonstrated (Figure 5.6). The 5 mM DFMO decreased *R. leguminosarum* growth, but the exogenous addition of 0.5 mM spermidine, norspermidine or homospermidine restored growth to nearly wild-type levels. The homo-spermidine complementation indicated that the DFMO growth effect is due to triamine depletion and not a side-effect due to growing the cells with a potentially toxic compound. The ability of all three triamines to increase growth indicated that the exact structure of the triamine is not critical for growth. It is interesting that a non-native triamine was able to restore growth fully, rather than being toxic to the cell. *R. leguminosarum* has evolved to contain homospermidine as its triamine, and whatever role it is performing within the cell must be able to be wholly replaced by spermidine or norspermidine. Spermidine is one carbon shorter, and norspermidine is two carbons shorter than the native triamine, homospermidine (Figure 3.2). It appears that all three exogenous triamines were able to perform the functions normally performed by intracellular homospermidine within the cell. It is highly likely that the intracellular polyamine-interacting structures are adapted for accepting a homospermidine molecule, rather than a norspermidine or spermidine. It appears that there is flexibility in structure, and the exact length and charge of the triamine in *R. leguminosarum* are not crucial to the function.

It was demonstrated that *R. leguminosarum* was able to uptake any (native and non-native) exogenous polyamine supplied (detected by HPLC of intracellular extracts, Figures 5.5 and 5.7). This suggests that the uptake systems in *R. leguminosarum* must be dynamic. The idea that *R. leguminosarum* has developed and evolved to produce and use homospermidine as its triamine, but when necessary is able to use alternatives, norspermidine or spermidine, is fascinating. Although homospermidine, norspermidine and spermidine are similar, as they are all positive charged triamines, they have slightly different length carbon chains (Figure 3.2). The role of polyamines in *R. leguminosarum* is relatively unstudied, but if we assume that they are performing similar roles in *R. leguminosarum* as in other documented species, they may be binding with proteins and membrane-bound systems and intercalating with DNA. So, the principle that an organism normally synthesises one triamine, but is able to use alternative triamines, suggests that the bacteria are able to be biochemically flexible in order to survive. It can be hypothesised that carbon length is an important feature of the biochemical interactions polyamines perform, such as intercalation with DNA. However, the systems must be more adaptable than previously thought.

A possible area for further study would be to investigate the phenotypic changes in the *R. leguminosarum* cells with a non-native triamine. The *R. leguminosarum* cells are clearly able to grow, but investigations into their ability to fix nitrogen and remain motile are all possible areas for investigation. This information may explain definitively the role of homospermidine, and other polyamines, in *R. leguminosarum.* In addition to spermidine and norspermidine, other triamines and polyamine analogues such as aminobutylcadaverine, carboxynorspermidine and carboxyspermidine could be tested for their ability to restore the *R. leguminosarum* growth defect caused by DFMO.

It was discovered that a three-fold decrease in the sole triamine (homospermidine) in *R. leguminosarum*, does not have a detectable growth or physical phenotype. These data generated in this investigation showed that when a transposon (Tn5) was inserted into the *R. leguminosarum* genome 100 bp up-stream of the translational start codon of the *hss*, homospermidine was

still present intracellularly, but the concentration was three-fold less when compared to the wild-type. This suggests that the organism lives with an excess of polyamine under normal physiological conditions. It can be hypothesised that the large pool of homospermidine is necessary for survival under stressful conditions. The exact role of homospermidine in Rhizobium species has not been reported upon. There are theories which link the presence to homospermidine to nitrogen fixation [23], but this hypothesis has yet to be proved.

C. jejuni

The other bacterial model used in this investigation was *C. jejuni*. The polyamine profile of the micro-organism was demonstrated for the first time using HPLC (Figure 5.11) and the possible biosynthetic pathway of the polyamines in *C. jejuni* was described (Figure 3.4). Genetic manipulation of the polyamine pathway by insertion inactivation of the carboxyspermidine decarboxylase (CANSDC) gene in *C. jejuni* resulted in cells with a severely reduced growth phenotype. This phenotype could be restored in two ways; firstly by genetic complementation with the disrupted gene, and secondly with the addition of exogenous polyamines (Figures 5.12 and 5.13). CANSDC was predicted to form spermidine by the decarboxylation of carboxyspermidine. The polyamine profile of *C. jejuni* was established as being almost completely the triamine spermidine, as no other polyamines were detected using the HPLC method (Figure 5.11). This information agreed with the predicted polyamine biosynthetic pathway (Figure 3.4) based on sequence homology with previously annotated polyamine biosynthetic orthologues. However, no putrescine was seen intracellarly in *C. jejuni*. It is possible that the putrescine was there, but at levels not detected by this method, or that the putrescine was quickly metabolised into carboxyspermidine, and therefore not seen when investigated by HPLC analysis of cell extracts. An area of possible investigation regarding the polyamine profile of *C. jejuni*, would be to further test the intracellular polyamines present over a time course, as the polyamine profile may change when the cells are in exponential or stationary growth phases.

The *C. jejuni* mutant defective in CANSDC was shown, by HPLC analysis, to lack its native triamine, spermidine. The CANSDC mutant had a severely reduced growth phenotype, but this phenotype could be partially restored with the addition of exogenous native and non-native triamines; spermidine, homospermidine and norspermidine (Figure 5.12). Norspermidine is one carbon shorter than spermidine, and spermidine is one carbon shorter than homospermidine (Figure 3.2). The ability of all three to restore *C. jejuni* growth to a similar level suggested that whatever role spermidine (the native polyamine) is fulfilling intracellularly, it can be restored by a structurally similar polyamine. The cellular interactions and functions of the triamine in *C. jejuni* are unknown. However, this ability to use alternative polyamines suggested that the exact charge, mass or length of the polyamine is not crucial for growth.

The ability of all the triamines to appear intracellularly in *C. jejuni* suggested that the cells have transporters, or transmembrane channels, which are allow the passage of non-native polyamines into the cell.

The phenomenon regarding triamine specificity in *C. jejuni* was further investigated by complementing the CANSDC mutant, which is triamine deficient, with a non-native triamine-synthetic gene. The gene chosen was *Paramecium tetraurelia* homospermidine synthase (PARA-HSS), under a native Campylobacter promoter (*met*K). The aim was to uncover, using a genetically engineered organism, if the native triamine could be replaced with another by *in vivo* synthesis in *C. jejuni*. Unfortunately, the genetic complementation of the CANSDC mutant (which lacked any triamine) with PARA-HSS was unable to synthesise an alternative triamine, homospermidine. PARA-HSS ordinarily uses two molecules of putrescine to form one of homospermidine, based on HPLC analysis of polyamines formed in an *E. coli* host expressing PARA-HSS (Figure 4.6). However, the possibility that the intracellular environment of *C. jejuni* is not suitable for the PARA-HSS to function cannot be ruled out. If there was no putrescine in *C. jejuni*, then an improved way to develop this experiment would be to genetically manipulate the *C. jejuni* so that both the CANSDC and the preceding biosynthetic gene in the predicted pathway, carboxyspermidine dehydrogenase (CANSDH), are inactive (Figure 3.4). Then, HPLC of cell extracts from the double mutant (CANSDC & CANSDH)

would confirm a build up of putrescine, if the predicted pathway is correct. This double-mutant could then be manipulated to express the PARA-HSS gene. Alternatively, a single insertion mutation in CANSDH would have a similar effect (no carboxyspermidine production) and could be used in the same way. Making a single mutant (and complementing it with a single gene) is technically and experimentally easier than a double mutant when the genes are non-consecutive. This experiment would allow understanding of whether *C. jejuni* is able to synthesise non-native polyamine *in vivo*. This is something which has never been investigated before.

It is unknown if Figure 3.4 is the correct pathway for polyamine biosynthesis in *C. jejuni*, because the biochemical validation of the genes has not been performed. This investigation assumed that the gene annotation and predicted polyamine pathway was correct.

A simple way to demonstrate the activity of PARA-HSS, when it is expressed in CANSDC-inactivated *C. jejuni,* would be to grow the cells in a medium supplemented with putrescine. Although it has not been shown that *C. jejuni* has the ability to take up putrescine, the bacterium is able to take up a variety of different polyamines, therefore it is likely it can move putrescine into the cytoplasm. This would be a logical experiment to perform. It can be hypothesised that if putrescine could enter the cells, then the PARA-HSS which is found intracellularly could use two molecules of putrescine to form a homospermidine molecule. In addition to the lack of substrate, another possible explanation for the lack of intracellular PARA-HSS activity could be due to co-factor availability. There may be a co-factor, which is yet to be reported, crucial for PARA-HSS to function and is available in *E. coli,* but not *C. jejuni.* This investigation did not use the proteomic approach to determine if the PARA-HSS enzyme was being expressed in the *C. jejuni* cells. It cannot effectively be determined without experimentation whether the PARA-HSS protein would be seen on the gel, or whether it would masked by the native *C. jejuni* proteins with a similar mass and isoelectric focusing point.

The exact role of polyamines in *C. jejuni* is yet to be understood, but further phenotypic studies using the wild-type strain, and the CANSDC mutant, are options for the future. The only phenotype investigated in the CANSDC

mutant was growth, however there is a suite of *C. jejuni* phenotype experiments which can be performed. The ability of the mutant, and genetically complemented mutant, to swarm and swim, form biofilms and survive stressful conditions could all be tested and compared with the wild-type. In addition, microarrays defining the regulation of genes within the mutant compared to the wild-type and proteomics to investigate the presence or absence of particular proteins are both options for further investigation in this field.

5.5 Concluding Remarks

This investigation has shown using two distinct bacteria model organisms that each contain specific polyamine profile that is necessary for optimum growth. Using both chemical and genetic manipulation, the polyamine levels have been depleted in these organisms, and this in turn had a detrimental effect on growth. However, it has been shown that both micro-organisms were able to take up non-native polyamines and growth was returned to near wild-type levels. This raises interesting questions regarding why a micro-organism has evolved to contain one particular polyamine when other polyamines, significantly different in length to the natural polyamine, are able to perform the same function. It is possible that this promiscuity of polyamine use occurred because of environmental pressures, however, it has been shown that polyamines are a basic requirement for a normal phenotype. Although the exact role of polyamines in these model organisms remains poorly defined, the conclusions formed in this investigation will help in the understanding of polyamine metabolism.

Chapter Five – References

1. Kusano T, Berberich T, Tateda C, Takahashi Y: **Polyamines: essential factors for growth and survival**. *Planta* 2008, **228**(3):367-381.
2. Busse HJ, Schumann P: **Polyamine profiles within genera of the class Actinobacteria with LL-diaminopimelic acid in the peptidoglycan**. *International Journal of Systematic and Evolutionary Microbiology* 1999, **49 Pt 1**:179-184.
3. Forouhar F, Lew S, Seetharaman J, Xiao R, Acton TB, Montelione GT, Tong L: **Structures of bacterial biosynthetic arginine decarboxylases**. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* 2010, **66**(12):1562-1566.
4. Ding H, Hynes MF: **Plasmid transfer systems in the rhizobia**. *Canadian Journal of Microbiogy* 2009, **55**(8):917-927.
5. Glian'ko AK, Akimova GP, Makarova LE, Sokolova MG, Vasil'eva GG: **Oxidation processes at initial stages of interaction of root nodule bacteria (*Rhizobium leguminosarum*) and pea (*Pisum sativum L.*): a review**. *Prikl Biokhim Mikrobiol* 2007, **43**(5):576-582.
6. Fujihara S, Harada Y: **Fast-growing root nodule bacteria produce a novel polyamine, aminobutylhomospermidine**. *Biochemical and Biophysical Research Communications* 1989, **165**(2):659-666.
7. Pearson BM, Gaskin DJ, Segers RP, Wells JM, Nuijten PJ, van Vliet AH: **The complete genome sequence of *Campylobacter jejuni* strain 81116 (NCTC11828)**. *Journal of Bacteriology* 2007, **189**(22):8402-8403.
8. Young JP, Crossman LC, Johnston AW, Thomson NR, Ghazoui ZF, Hull KH, Wexler M, Curson AR, Todd JD, Poole PS *et al*: **The genome of *Rhizobium leguminosarum* has recognizable core and accessory components**. *Genome Biology* 2006, **7**(4):R34.
9. Boyle SM, MacIntyre MF, Sells BH: **Polyamine levels in *Escherichia coli* during nutritional shiftup and exponential growth**. *Biochimica et Biophysica Acta* 1977, **477**(3):221-227.
10. Yamamoto S, Nagata S, Kusaba K: **Purification and Characterization of Homospermidine Synthase in *Acinetobacter tartarogenes Atcc-31105***. *Journal of Biochemistry* 1993, **114**(1):45-49.

11. Böttcher F, Ober D, Hartmann T: **Biosynthesis of pyrrolizidine alkaloids: putrescine and spermidine are essential substrates of enzymatic homospermidine formation**. *Canadian Journal of Chemistry* 1994, **72**(1):80-85.
12. Fox MA, Karunakaran R, Leonard ME, Mouhsine B, Williams A, East AK, Downie JA, Poole PS: **Characterization of the quaternary amine transporters of *Rhizobium leguminosarum* bv. viciae 3841**. *FEMS Microbiology Letters* 2008, **287**(2):212-220.
13. Hamana K, Minamisawa K, Matsuzaki S: **Polyamines in Rhizobium, Bradyrhizobium, Azorhizobium and Argobacterium**. *FEMS Microbiology Letters* 1990, **71**(1-2):71-76.
14. Braeken K, Daniels R, Vos K, Fauvart M, Bachaspatimayum D, Vanderleyden J, Michiels J: **Genetic determinants of swarming in *Rhizobium etli***. *Microbial Ecology* 2008, **55**(1):54-64.
15. Lee MJ, Huang CY, Sun YJ, Huang H: **Cloning and characterization of spermidine synthase and its implication in polyamine biosynthesis in *Helicobacter pylori* strain 26695**. *Protein Expression and Purification* 2005, **43**(2):140-148.
16. Hamana K, Saito T, Okada M, Niitsu M: **Polyamine Distribution Profiles among Some Members within Delta-and Epsilon-Subclasses of Proteobacteria**. *Microbial Culture Collection* 2004, **20**(1):3-8.
17. van Vliet AH, Baillon MA, Penn CW, Ketley JM: **The iron-induced ferredoxin FdxA of *Campylobacter jejuni* is involved in aerotolerance**. *FEMS Microbiology Letters* 2001, **196**(2):189-193.
18. Wosten MM, Boeve M, Koot MG, van Nuenen AC, van der Zeijst BA: **Identification of *Campylobacter jejuni* promoter sequences**. *Journal of Bacteriology* 1998, **180**(3):594-599.
19. Wallace HM, Fraser AV: **Inhibitors of polyamine metabolism: review article**. *Amino Acids* 2004, **26**(4):353-365.
20. Bacchi CJ, Yarlett N: **Effects of antagonists of polyamine metabolism on African trypanosomes**. *Acta Trop* 1993, **54**(3-4):225-236.
21. Kallio A, McCann PP: **Difluoromethylornithine irreversibly inactivates ornithine decarboxylase of *Pseudomonas aeruginosa*, but does not inhibit the enzymes of *Escherichia coli***. *Biochemistry Journal* 1981, **200**(1):69-75.

22. Barry DP, Asim M, Leiman DA, de Sablet T, Singh K, Casero RA, Chaturvedi R, Wilson KT: **Difluoromethylornithine Is a Novel Inhibitor of Helicobacter pylori Growth, CagA Translocation, and Interleukin-8 Induction**. *PLoS One* 2011, **6**(2):e17510.

23. Hamana K, Miyagawa K, Matsuzaki S: **Occurrence of Sym-Homospermidine as the Major Polyamine in Nitrogen-Fixing Cyanobacteria**. *Biochemical and Biophysical Research Communications* 1983, **112**(2):606-613.

Chapter Six

Polyamines: Concluding Summary and Future Directions

Chapter Six - Conclusions of Polyamine Section

Physiology, necessity and specificity of polyamines

This investigation has provided some new and exciting insights into the necessity and specify of polyamines in bacteria. The absolute requirement for polyamines in bacteria was demonstrated using two different bacterial models; representatives from the α- and ε-Proteobacteria, *Rhizobium leguminosarum* and *Campylobacter jejuni* respectively. Both showed a decreased growth phenotype, compared to the wild-type, when they possessed a reduced polyamine profile. The polyamines in *R. leguminosarum* were almost completely removed by growing the cells in DFMO-supplemented media, a polyamine biosynthetic enzyme inhibitor. Genetic inactivation of one of the polyamine biosynthetic genes in *C. jejuni* led to a decreased polyamine content. In both species, this defective growth phenotype could be restored by the addition of exogeneous polyamines, in particular triamines. This suggested that both micro-organisms have the ability to take up these molecules, and must have transporters or transmembrane channels suitable for this purpose. Polyamine uptake in *R. leguminosarum* and *C. jejuni* has not been reported previously. In addition, it has been shown that the exact length of the triamine is unimportant in respect to the restoration of the phenotype. This is the first investigation regarding triamine supplementation to bacterial cells. These species have evolved to contain only one triamine, but can actually function and use a variety of non-native triamines when necessary. It can be hypothesised that under environmental conditions, these species may be able to use whichever triamine is available, and when there are no trimaines present, they are stimulated into producing their own.

One favoured concept is that the polyamine pathways have evolved and are distinct from one another, maybe circumstantial, rather than a necessary, evolutionary development. It can be hypothesised that the resulting polyamine pathways present in *R. leguminosarum* and *C. jejuni* due to evolutionary divergence have come to be more by chance, than by specific, directed evolution. It is known that several different triamines are able to function in both *R. leguminosarum* and *C. jejuni,* so it is possible that the

divergence of these species' polyamine synthetic profiles happened by chance, as the exact nature of the triamine is not a specific requirement.

Evolution of bacterial biosynthetic pathways is an interesting concept and it is an established fact that during bacterial evolution, redundant genes have been lost and new genes have been acquired by various species. In addition, genes can move between species by horizontal gene transfer, and this phenomenon is especially notable in single-cell organisms, such as bacteria. Evolution is far from perfect design. It is believed that a random, spontaneous mutation can be retained over several generations and lineage splits to become an area of the genome conserved across many species. It is possible, that the promiscuity of both *R. leguminosarum* and *C. jejuni* to use any of the triamines tested in this investigation, may be an example of the sporadic and undirected nature of evolution.

The essential nature of polyamines has been described in this thesis and there is particular interest in understanding which genes are fundamental for life. The results discussed here may add valuable knowledge to the investigations into genome design, where the objective is to design a bacterial cell with a completely synthetic genome. It is essential in genome design investigations that the genes which are shown to be essential for life are included. As this work develops and continues, a more species defined approach may be taken. And under these circumstances, it may be necessary for the representatives of the α- and ε-Proteobacteria to be analysed. The work presented in this thesis may be the backbone for future investigations into the essential nature of specific bacterial genes.

The results and discoveries in this investigation are of particular interest to those with an interest in evolutionary biology. The concept of evolution is not fully understood, and elucidation of this theory, using specific examples such as those discussed here, will help with the wider understanding of bacterial divergence.

The polyamine synthesis pathway in Campylobacter has not been fully characterised

A possible polyamine biosynthetic pathway has been predicted based on the sequence similarity of the *C. jejuni* genes with previously characterised or annotated genes (Figure 3.2). However, without robust biochemical testing of each of these genes, it is impossible to determine whether the predicted pathway is correct. The polyamine profile of *C. jejuni* is relatively simple, and HPLC analysis indicated that the cells only contained one polyamine: spermidine (Figure 5.11). However, the investigation to determine if the natural triamine in *C. jejuni* could be replaced by an alternative non-native triamine using genetic manipulation was unsuccessful. The fascination of the polyamine synthetic pathway in *C. jejuni* continues, and there are several areas in which to progress this investigation to determine more specifically about polyamine biosynthesis in *C. jejuni*.

Using Campylobacter as a model organism and gaining insights into its cellular functions has interesting implications in the fields of food safety and human infection. Campylobacter is a human and animal pathogen and has been identified as one of the leading causes of food poisoning in the developed world [1, 2]. Complications arising from campylobacter infection include serious and life-threatening syndromes, such as Guillain-Barré syndrome. Therefore, any discoveries found in the metabolism, survival or phenotype *C. jejuni* during this investigation, may have interesting implications in the pathogenesis of the organism. Discoveries from the continuation of this investigation could have applications in medical research and food safety.

Phylogenetic distribution of homospermidine synthetic genes

The range of organisms identified as either synthesising homospermidine or containing an *hss* gene is diverse. Figure 3.7 depicted the 'The Tree of Life' and clearly showed examples of organisms from the Bacteria, Archaea and Eukaryotes all containing *hss* orthologues. As our ability to sequence genomes becomes quicker, cheaper and easier, more genetic information is being

identified from more unusual species. Even unculturable organisms which are still undefined taxonomically are able to be sequenced using a metagenomic approach.

Looking at the phylogenetic distribution of the *hss* genes, it has been predicted that the *hss* in the α-Proteobacteria originated by vertical gene transfer. As the gene is found in most of the sequenced α-Proteobacteria species, it can be hypothesised that the gene was also present in the last common ancestor, and has been conserved throughout evolution. The presence of *hss* in the more phylogenetically diverse species can be attributed to horizontal gene transfer. The sporadic distribution of orthologues is a characteristic feature of horizontal gene transfer. Some of the specifics regarding of *hss* distribution were discussed in Section 4.4.2 and have been reported by Shaw *et al.* [3].

Functional genomic investigations, such as this, can be applied to a number of different biosynthetic or metabolic pathways. If the correct expression host and favourable conditions are used, it is possible that a number of investigations to validate gene function are possible. The extrapolation of this method has interesting implications for bioinformaticians and geneticists. The more biochemical validation of specific genes available in published databases, the more robust the bioinformatic analysis becomes. Validation of genome annotation is essential and this method has nearly limitless applications for future research.

Further insights into the homospermidine synthase (HSS) enzyme

The activity of the α-Proteobacterium *Bradyrhizobium japonicum* HSS was confirmed to possess kinetic parameters similar to that of previously characterised HSS proteins. This was the first study on a recombinantly expressed and purified HSS, all previous characterisations had been performed on cell extracts from species which naturally express HSS. The HSS activity of six previously functionally undefined genes was confirmed by expression in *E. coli*, followed by HPLC analysis of the cell extracts. The six *hss* genes were originally identified based on their sequence homology to other

putatively annotated *hss* genes. This was the first example of several *hss* genes, from phylogenetically diverse organisms, expressing homospermidine in *E. coli.*

This investigation has been a valuable phylogenetic tool. It has been shown, that the activity of a gene can be defined even though the organism it originates from is unculturable in the laboratory. The results from this investigation have provided an insight into polyamine metabolism in unculturable organisms and those which are difficult to grow in the laboratory. The conclusions drawn from this work suggest that bacteria, specifically *E. coli*, can synthesise non-native polyamines when expressing the necessary biosynthetic enzymes. In addition, *E. coli* cells are able to survive with these additional, unnatural polyamines.

The promiscuous substrate use of HSS has been demonstrated. When the six *hss* were expressed in *E. coli*, additional polyamine products were detected, such as diaminopropane and aminobutylcadaverine. This is the first documented example of the production of aminobutylcadaverine during *in vivo hss* expression. This conclusion raises interesting questions into why and how the enzymes have evolved to have the ability to make several different polyamines, and whether this information can be used in the future in research areas such as drug design, and functional genomics.

Further insights into the substrate promiscuity of the various HSS could be determined by looking at the purified proteins, or by investigating the crystal structures of the enzymes. It can be hypothesised that there may be specific attributes in the enzyme active site which enable multiple substrates to be metabolised, and these characteristics could be used in further enzymological investigations in polyamine synthetic proteins and beyond.

An interesting perspective is that when homospermidine (and HSS) were first discovered in the 1970s, the polyamine was classified as ‘uncommon’ because it had never been identified before. Now, *hss* orthologues are confirmed from the metagenomic studies of samples from our oceans, and it cannot be ruled out that homospermidine may be one of the most abundant polyamines on the

planet. This information and technology is relatively new, and the understanding of the subject of metagenomics will grow considerably over the next few years.

Future Directions

There are several areas to continue this investigation. Firstly, it should be noted that there are readily available phenotypic investigations which can be performed on the model organisms *R. leguminosarum* and *C. jejuni* after polyamine depletion. Further investigation using these model organisms could allow further understanding about the essential nature of polyamines in bacteria. *R. leguminosarum* is an excellent model organism for further study because, even without a detectable polyamine complement, it is still viable. There are reports of the micro-organisms lacking polyamines being unable to survive, or being debilitated to an extent which makes investigations difficult, reviewed by Cohen [4]. Results from these phenotypic investigations could provide answers regarding the specific function of polyamines in *R. leguminosarum* and *C. jejuni*. Some precise future directions are detailed below:

- Phenotypic investigations of *R. leguminosarum* including swimming and swarming ability, nitrogen fixation, and light, or electron, microscopy would provide information about the individual cell morphology. Cellular robustness could be investigated using sustainability testing, such as shocking the cells with extreme temperatures, pH or dryness.

- Gene regulation in polyamine depleted *R. leguminosarum* could be tested by microarray analysis. It could be hypothesised that to overcome the depleted polyamine content, the polyamine biosynthetic genes would be up-regulated by a compensatory mechanism. However, there may be more biochemical pathways which are altered in order for the *R. leguminosarum* to compensate for a reduced polyamine complement. The number of studies combining polyamine profile and microarray analysis are limited [5, 6], especially in micro-

organism models. A microarray such as this would be the first of its kind and provide further insight into the regulation and function of polyamines in *R. leguminosarum.*

- It would be novel to demonstrate the synthesis of a non-native triamine in *C. jejuni*, as this has never been shown before. This could be done by continuation of the genetic manipulation of the *C. jejuni* polyamine pathway, specifics for this investigation have been discussed in Section 5.4 : *C. jejuni*.

- Phenotypic studies of polyamine depleted *C. jejuni* such as motility, biofilm formation and survival in stressful conditions. In addition, the regulation of the genome and proteome in polyamine depleted *C. jejuni* could also be investigated using transcriptomic and proteomic approaches.

Chapter Six - References

1. DEFRA: **Zoonoses Report**. 2008, **PB13386**.
2. EFSA: **Scientific Opinion on Quantification of the risk posed by broiler meat to human campylobacteriosis in the EU**. *EFSA Journal* 2010, **8**(1).
3. Shaw FL, Elliott KA, Kinch LN, Fuell C, Phillips MA, Michael AJ: **Evolution and multifarious horizontal transfer of an alternative biosynthetic pathway for the alternative polyamine sym-homospermidine**. *Journal of Biological Chemistry* 2010, **285**(19):14711-14723.
4. Cohen S: **A guide to the Polyamines**. New York: Oxford University Press; 1998.
5. Chattopadhyay MK, Chen W, Poy G, Cam M, Stiles D, Tabor H: **Microarray studies on the genes responsive to the addition of spermidine or spermine to a *Saccharomyces cerevisiae* spermidine synthase mutant**. *Yeast* 2009, **26**(10):531-544.
6. Igarashi K, Kashiwagi K: **Polyamine Modulon in *Escherichia coli*: genes involved in the stimulation of cell growth by polyamines**. *Journal of Biochemistry* 2006, **139**(1):11-16.

Chapter Seven

Introduction to Campylobacter and Selenoproteins

TABLE OF FIGURES

Chapter Seven - Introduction to Campylobacter and Selenoproteins

7.1 *Campylobacter jejuni*

7.1.1 Introduction

The Campylobacter genus, which means "curved bacteria," was classified in 1963 by Sebald and Véron [1], although Campylobacter was documented in 1880 by Theodore Escherich. The genus Campylobacter is a member of the Campylobacteraceae family, within the ε-Proteobacteria sub class. The Campylobacteraceae includes a large group of organisms which colonise hostile environments, like the gastro-intestinal tract of humans and animals. The Campylobacter genus includes both pathogenic and commensal organisms, and includes 18 species and subspecies [2]. More specifically, and for the scope of this project, the focus will be on *Campylobacter jejuni subspecies jejuni*. To date, 13 genomes of *C. jejuni* strains have been the sequenced and published.

The Campylobacter genus contains Gram-negative, spiral, microaerophillic micro-organisms which are capable of causing of food poisoning in the developed world. Infection of humans with campylobacter results in diarrhoea, but occasionally systemic infection occurs; the resulting illness can be dehabilitating and life-threatening [3, 4]. Several features about *C. jejuni* survival, colonisation in a host, metabolism, phenotype and genome are known [5-8]. However, some aspects of *C. jejuni* biology and virulence remain undefined, and this investigation will study the organism with specific genes inactivated to determine their role in *C. jejuni* biology. It is necessary to understand the exact mechanisms used by *C. jejuni*, as is it not possible to extrapolate information gained from other bacteria, such as *Salmonella* spp. or *E. coli.* Molecular biology and biochemistry will be used together to learn more about this human pathogen. *C. jejuni* has a relatively small genome (1.7 Mb), but this does not make it any less complex to understand. The whole *C. jejuni* genome will be investigated in this study using a transcriptome approach. This pathogen is able to continually elude the host's immune

system and survive in a variety of environments. This makes *C. jejuni* an important organism for further study and investigation.

7.1.2 Epidemiology of Campylobacter

The relationship between Campylobacter and human illness was first clearly demonstrated in 1972, and, according to the DEFRA Zoonoses Report, Campylobacter is the most commonly reported gastrointestinal pathogen in the UK. Estimates suggest that the total annual number (including unreported) of campylobacteriosis cases could be as high as 445,000 [3]. In the 27 countries of the EU, it has been estimated that between two and 20 million cases occur annually [4]. The thermophilic Campylobacter species *C. jejuni* and *C. coli* are the most common causes of human campylobacter infection. In 2008, of the 405 UK cases speciated, 96% were *C. jejuni* while the remaining 4% were *C. coli* isolates [3].

Most cases of Campylobacter infection are sporadic and the route of transmission is usually from the consumption of under-cooked meat, or cooked meat which has been cross-contaminated with raw meat. Raw poultry is often contaminated because *C. jejuni* colonise the deep crypts of the chicken caecum at high levels (10^{10} colony-forming units per gram of colonised intestine) [8], and the caecal and feacal content can contaminate the meat during the slaughter process [9]. The infective dose of *C. jejuni* is considered to be small [10]; human feeding studies suggest that about 500-800 bacteria may cause illness in some individuals [11, 12]. The documented rate of campylobacter infection varies between developed countries; this could be due to several factors, including food manufacturing standards, food choice, bacterial levels in the animals and the reporting and testing procedures. In addition, the rate of infection fluctuates seasonally and geographically [13]. In the UK, the incidence of infection with Salmonella appears to be steadily falling (Figure 7.1), this is most likely due to the vaccination of chickens against Salmonella [14]. The infection incidence of Campylobacter cases is not decreasing and remains higher than the infection incidence of Salmonella. These data indicate the importance of pathogenesis studies on the Campylobacter species.

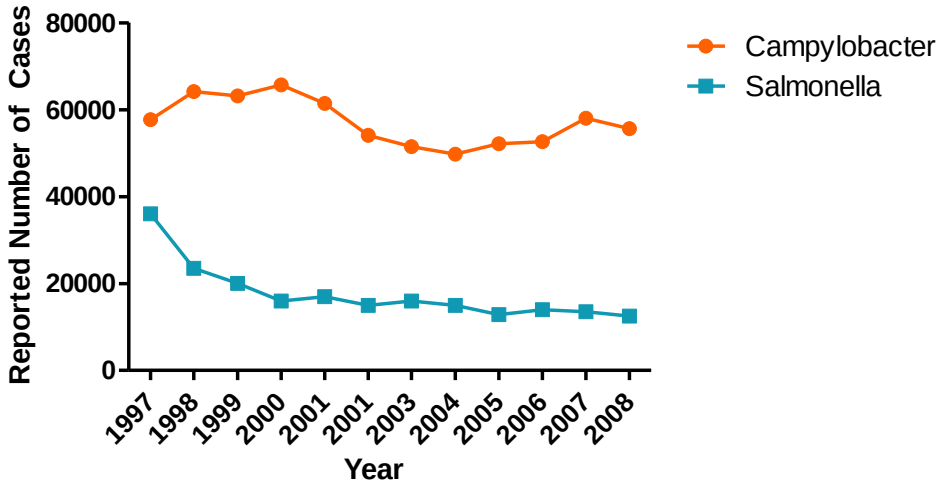


Figure 7.1 Campylobacter and Salmonella infection incidences
(Data taken from Public Health Laboratories UK figures)

A human infection by Campylobacter results in inflammatory, sometimes bloody, diarrhoea or dysentery syndrome, most cases include cramps, fever and pain [10]. Although campylobacteriosis is generally self-limiting, any gastro-intestinal disease has an important economic impact in the developed world. Campylobacter is hyperendemic in developing countries, owing to poor sanitation and close human contact with animals [15].

Serious complications can arise following a Campylobacter infection, such as Guillain-Barré Syndrome (GBS) and Miller-Fisher Syndrome, which represent a series of neurodegenerative paralytic disorders associated with degeneration of motor neurones. Between one and two cases of GBS occur per 100,000 people in the UK and USA [16]. This auto-immune disorder of the peripheral nervous system occurs several months after *C. jejuni* infection and is the most common cause acute flaccid paralysis. Campylobacter-associated reactive arthritis is more common following Campylobacter infection, with an annual incidence of 4.3 per 100,000 [17].

7.1.3 Campylobacter Biology

Most species of Campylobacter have either uni- or bi-polar flagella and a curved appearance, although S-, V- and comma-shaped campylobacter are found. The Campylobacter cells are normally free-floating, but occasionally

observed in long chains. Campylobacter species are oxidase-positive, catalase-positive and are commonly urease-negative, although urease-positive thermophilic Campylobacter species have been reported [18]. These microaerophilic and capnophilic organisms grow optimally in an atmosphere containing 5 - 10% oxygen, 5 - 15% carbon dioxide and a temperature between 34 - 44°C [19]. Unlike *Escherichia coli*, Campylobacter require a specific complex growth medium, since they lack many of the genes needed to degrade carbohydrates. *C. jejuni* neither oxidises or ferments glucose, but is capable of amino acid catabolism [20]. In addition, *C. jejuni* has been shown to be capable of utilising a variety of electron donors and acceptors which are used in the electron transport chain to support *C. jejuni* growth [20, 21]

7.1.4 Campylobacter Pathogenesis & Virulence

The exact mechanism by which *C. jejuni* causes illness is yet to be fully understood. The contact of *C. jejuni* with human epithelial cells and invasion of the epithelial cells, are triggers for differential *C. jejuni* gene expression [22, 23]. Research into the methods by which *C. jejuni* survive and cause disease in humans is on-going, but unfortunately, a robust animal model for studying the pathogenic effect of Campylobacter is not readily available.

It is essential that during *C. jejuni* pathogenesis the bacteria adhere to the host. This is not fully understood in *C. jejuni* and most likely is a multifactorial process, with the flagella holding an important role [5]. The *C. jejuni* then invades the host cells and the infection takes place [24]. *C. jejuni* has a number of survival mechanisms, such as the flagella (for rapid movement), biofilm production, lipooligosaccharide and capsular polysaccharide formation, toxin production and acid-resistance. These virulence factors, along with other mechanisms, are used to evade the host immune responses and continue infection [7, 8].

In order to infect a human host, it is essential that *C. jejuni* can withstand the gastric conditions (approximately pH 2 - 3). The optimum pH for *C. jejuni* growth is 6.5 - 7.5, and the organism does not grow well below pH 4.9 or above 9.0 [25]. It is possible that *C. jejuni* survives the severe gastric pH by being encased in the food item they originated from, and additionally, under a

traditional Western diet, meal ingestion will increase the median gastric pH to about 6 [26]. However, it is likely that *C. jejuni* still require mechanisms for survival at low pH. Survival at acidic pH values is temperature dependant, but cellular inactivation is rapid at pH values less than 4.0, especially above refrigeration temperatures [27]. Studies on the subject suggest that acid tolerance may be strain-dependent, and Murphy *et al.* [28] have reported upon adaptive acid tolerance. This phenomenon occurs when the bacteria are exposed to sub-lethal acid and aerobic conditions. Reid *et al.* [29, 30] investigated the genes responsible in *C. jejuni* for surviving acid-shock and highlighted the important role played by cell surface components (flagella, the outer membrane, capsular polysaccharides and lipooligosaccharides). The down-regulation of genes encoding ribosomal proteins was observed and this suggested the reshuffling of energy toward the expression of components required for survival was taking place. Acid-shock also caused *C. jejuni* to up-regulate genes involved in stress responses, such as heat shock genes and oxidative and nitrosative stress response genes [29, 30].

7.1.5 Campylobacter Laboratory Techniques

Using a specific growth medium and microaerobic conditions, Campylobacter is successfully grown and used experimentally. The natural resistance to some antibiotics, together with antibiotic resistance gained by genetic manipulation, is used as a tool to screen between different *C. jejuni* strains. Most strains of *C. coli* and many strains of *C. jejuni* are naturally competent and can be transformed with naked DNA with no special treatment [31-33]. A shuttle plasmid is used which contains the antibiotic resistance cassette flanked by regions homologous to the target gene. Electroporation is also used to transform *C. jejuni* with plasmid DNA. The plasmid DNA homologously recombines into the *C. jejuni* genome and the target gene is inactivated. Simultaneously, the strain becomes resistant to the antibiotic.

Naturally, *C. jejuni* is resistant to several antibiotics including vancomycin, trimethoprim, streptogramin, rifampin B, novobiocin, bacitracin and usually cephalothin [34]. This antibiotic-resistant phenotype enables *C. jejuni* to be selected on a medium, such as Skirrow; an agar infused with a cocktail of antibiotics: vancomycin, polymixin-B and trimethoprim. Fluoroquinolone-

resistant campylobacters have been found in poultry faeces and carcasses, and in retail poultry meat products in most areas of the world. Normally, fluoroquinolones inhibit the growth of Campylobacter by binding to bacterial DNA gyrase and DNA topoisomerase IV. Routinely, fluoroquinolones are used to treat the animals from which meat products originate, this could explain the contamination of food with fluoroquinolone-resistant campylobacters [35].

C. jejuni can be genetically manipulated to express additional antibiotic resistance genes. Routinely, chloramphenicol, erythromycin and kanamycin are inserted into the *C. jejuni* genome to inactivate a target gene. Chloramphenicol is bacteriostatic; growth is stopped by inhibition of peptidyltransferase activity of the bacterial ribosome. The antibiotic molecule binds to A2451 and A2452 residues in the 23S rRNA of the 50S ribosomal subunit, preventing peptide bond formation. Kanamycin is bactericidal and from the aminoglycosides family. Bacterial death takes place because the aminoglycosides irreversibly bind to the 30S ribosome and freezes the 30S initiation complex (30S-mRNA-tRNA), so that no further initiation can occur. The aminoglycosides also slow down protein synthesis that has already initiated and induce misreading of the mRNA. Erythromycin is a macrolide and is bacteriostatic. The macrolides inhibit translocation of the peptidyl tRNA from the A to the P site on the ribosome by binding to the 50S ribosomal 23S RNA.

7.2 Campylobacter Genomics

C. jejuni NCTC 11168 has a circular chromosome of 1.7 Mbp (30.6% G+C) which is predicted to encode approximately 1,600 proteins and 54 stable RNA species [6, 36]. *C. jejuni* NCTC 11168 is the most gene-dense bacterium so far to be sequenced (completed in 2000), with 94.3% of the genome containing coding genes, according to the Sanger Institute. After re-annotation of the genome in 2007, *C. jejuni* NCTC 11168 is the most comprehensively annotated Campylobacter genome to date. Campylobacters have a relatively small genome, with only roughly 25–33% as many genes as *E. coli* strains [37]. Small genomes, such as *C. jejuni* NCTC 11168, would be expected to contain more essential genes and relatively fewer dispensable, or auxiliary, genes [38].

The *C. jejuni* genome is unusual in that there are virtually no insertion sequences or phage-associated sequences and very few repeat sequences. However, the *C. jejuni* genome contains several examples of hyper-variability. For example, the analysis of single-nucleotide polymorphisms (SNPs) in *C. jejuni* can discriminate between isolates [39]. Multi-locus sequence typing and a DNA microarray-based comparative genomic indexing approach have been used to examine the genomic diversity and gene content of *C. jejuni* subsp. *doylei* strains. The investigation identified subsets of absent genes and showed that these strains were phylogenetically distinct from *C. jejuni* subsp. *jejuni* strains [40]. Short homopolymeric runs of nucleotides were commonly found in genes encoding the biosynthesis or modification of surface structures, or in genes of unknown function [36]. Genomic diversity between *C. jejuni* isolates can possibly be attributed to these hyper-variable regions. It is possible that due to this feature, certain cells in one population may be able to withstand environmental pressures, such as acidity or temperature, and the population continues to survive post-stress. The apparently high rate of variation of these homopolymeric tracts may be important in the survival strategy of *C. jejuni* [36]. The amount of genomic variation between Campylobacter species has occurred via natural transformation, phase variation, plasmid transfer and infection with bacteriophages. This variation poses a continuous challenge for studies on pathogenesis, physiology, epidemiology and evolution of Campylobacter [41].

7.2.1 Campylobacter Gene Regulation and Expression

Understanding gene regulation contributes to the understanding of *C. jejuni* survival, growth and pathogenesis [15, 42]. Advances in molecular and computational biology have led to the development of powerful, high-throughput methods for the analysis of differential gene expression [43]. Whole transcriptome analysis, using microarray hybridisation, is readily available for *C. jejuni*. Microarray analysis can be used in conjunction with proteome profiles to gain global regulation understanding [42]. The up- and down- regulation of genes in *C. jejuni* can be investigated under normal and stressed conditions [29, 30]. However, unlike other organisms, the mechanisms of gene regulation are still poorly understood in *C. jejuni* [44].

To cope with stressful environments, micro-organisms often have genes which can be up- or down-regulated in response to the stress to aid survival. *C. jejuni*, like many other micro-organisms, are sensitive to environmental stresses, such as pH, oxygen and temperature. But interestingly, it lacks many of the homologues responsible for stress adaptation present in other gastrointestinal pathogens. For example, homologues of the sigma factors RpoH and RpoS, used for heat shock and entry into stationary phase respectively, are absent [45]. In addition, homologues of CspA and Lrp (proteins involved in temperature shock and metabolism regulation) are also missing from the *C. jejuni* genome. These facts make *C. jejuni* an interesting candidate for further gene regulation investigation.

Extreme temperatures, either excessive heat or cooling, can be detrimental to micro-organisms. Most have evolved to survive within a particular temperature range, but mechanisms are in place to aid survival when exposed to temperatures at the limit of survival. Konkel *et al.* [46] found that genes encoding 24 proteins were up-regulated when *C. jejuni* cells were heat-shocked at temperatures between 43 and 48°C. The proteins produced following heat-shock are known collectively as heat-shock proteins and are found in many bacterial species. Seventeen heat-shock proteins have been observed in *C. jejuni* [47], some are molecular chaperones [48] and others are linked to thermotolerance [46]. In another investigation using a whole genome microarray, Stintzi *et al.* [47] examined 1,626 *C. jejuni* genes and found 336 had altered expression in response to a temperature change.

It has been reported that a sudden downshift of the growth temperature triggers a drastic reprogramming of bacterial gene expression to allow cell survival under the new unfavorable conditions [49]. The cold-shock survival strategy studied in *E. coli* includes genes regulated at both transcriptional and post-transcriptional levels [49]. *C. jejuni* does not appear to produce cold-shock proteins, but even at temperatures as low as 4°C, vital cellular processes still function [50]. This suggests that *C. jejuni* has mechanisms in place for survival at cold temperatures. Cold-shock proteins have been indentified in other bacterial species, such as *Bacillus subtilis, Thermotoga maritima,*

Salmonella typhimurium and *E. coli*. They are typically small (67 - 73 amino acids) and consist of a single nucleic acid-binding cold shock domain [51]. Genes encoding energy metabolism and flagella assembly are up-regulated in *C. jejuni* when exposed to cold conditions [52]. This suggests that more energy is required by *C. jejuni* at lower temperatures.

Reid *et al.* [30] tested the gene expression of *C. jejuni* at several pHs and observed the down-regulation of genes encoding ribosomal proteins. It was concluded that this most likely reflects the need of *C. jejuni* to reshuffle energy toward the expression of biological components required for survival. The investigation also showed acid-shock caused *C. jejuni* to up-regulate genes involved in stress responses; including heat-shock genes and oxidative and nitrosative stress genes. Studies have eluded to the adaptation to acidic conditions by *C. jejuni* [28, 53]. This adaptation involves the differential expression of respiratory pathways, the induction of genes for phosphate transport, and the repression of energy generation and intermediary metabolism genes [29]. But the general lack of understanding of specific *C. jejuni* acid response mechanisms shows that this is clearly an area for further study.

7.3 Campylobacter Energy Metabolism

Thermotolerant campylobacters, such as *C. jejuni*, are the most frequent cause of bacterial infection of the lower intestine worldwide [54]. The human intestine is home to very large numbers of micro-organisms, with bacterial cells exceeding 1×10^{11} ml^{-1} in the colon [55]. Many commensal gastro-intestinal tract bacteria ferment dietary polysaccharides and other carbohydrates to a variety of organic acids, including acetate, propionate, butyrate, formate, succinate and lactate, which in turn can be used as carbon and electron sources for other bacteria, such as the pathogenic *C. jejuni* which cannot use organic sugars as energy sources [56].

The main carbon sources used by *C. jejuni in vivo* are likely to be amino acids, based on the fact that the bacterium is unable to metabolise exogenous sugars, as it does not contain the glycolytic enzyme 6-phosphofructokinase [57], but

does have the transport and enzymatic capacity for amino acid catabolism [21]. In *C. jejuni* NCTC 11168, genome sequence analysis together with consideration of the most likely catabolic pathways in which these predicted enzymes participate, predicts that *C. jejuni* completes the catabolism of aspartate, asparagine, glutamate, glutamine, serine and proline [20, 21]. Biochemical studies have been carried out on amino acid utilisation in *C. jejuni* [58, 59]. Both studies showed the rate of amino acid metabolism to be dependent on the concentration available. The rate of metabolism of glutamate, glutamine and alpha-ketoglutarate (produced by the deaminination of glutamate) by *C. jejuni* is greatly enhanced by increased substrate concentration [58]. *C. jejuni* growth is definitively linked to amino acid (serine, aspartate, glutamate and proline) catabolism. Slow growth in a complex medium shifted amino acid utilisation from more (serine and aspartate) to less preferred substrates (glutamate, proline and possibly amino acids from the proteolysis of peptones) [59].

Although it is known that *C. jejuni* is able to use lactate as both a carbon source and electron donor [56], less is known about formate. Unpublished data (Bruce Pearson and Ian Colquhoun, IFR) shows the ability of *C. jejuni* to metabolise formate after four hours growth (Figure 7.2). Westfall *et al.* [58] showed the constant, rapid production of CO_2 from formate during *C. jejuni* incubation (three hours). Formate respiration in *C. jejuni* has been reported upon during a study on the sulphite respiration system [60] and experimentally, *C. jejuni* has been shown to exhibit hydrogenase and formate dehydrogenase activities [61]. Respiratory activities of *C. jejuni*, determined with membrane vesicles, were 50- to 100-fold higher with formate and hydrogen than with succinate, lactate, malate or NADH as substrates. This result suggested that the metabolism of hydrogen and formate probably serves as the major sources of energy for *C. jejuni* growth [62] and formate is perhaps one of the preferred energy sources.

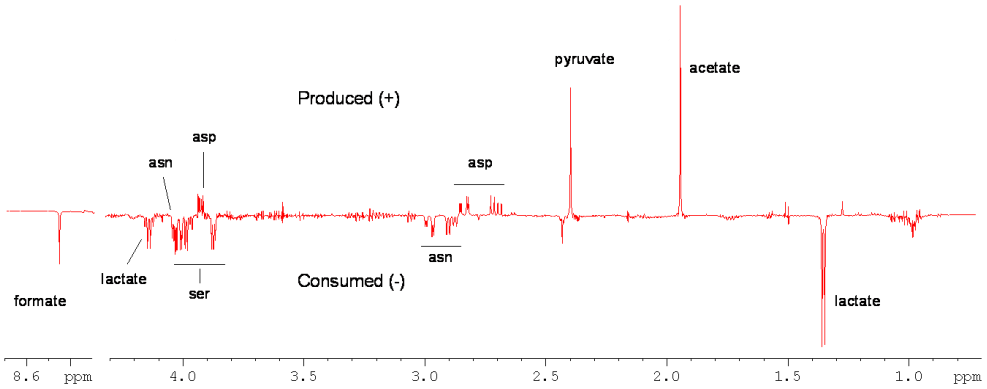


Figure 7.2 ^{1}H NMR of *C. jejuni* growth medium
Changes in the *C. jejuni* growth medium (Brucella) composition after four hours growth shown by comparison of differential spectra, ^{1}H NMR, asn - asparagine, asp – aspartate, ser – serine. The upward peaks represent the production (+) of a compound, the downward peaks represent the consumption (-) of a compound.

7.3.1 *Campylobacter jejuni* Formate Dehydrogenase (FDH)

Formate dehydrogenases (FDHs) are a set of enzymes that catalyse the oxidation of formate. During the reaction, electrons are donated to a second substrate, such as NAD+. NAD-dependent FDHs have been identified in a large number of both Gram-negative and Gram-positive bacterial species and, due to their simple mechanistic nature, have been used as a model for investigating catalysis mechanisms involving hydride ion transfer [63].

$$\mathbf{Formate + NAD^{+} + H^{+} + 2e^{-} \leftrightarrow CO_2 + NADH + 2H^{+} + 2e^{-}}$$

Formate is a very good electron donor, as the formate/bicarbonate couple has a highly negative redox potential (E_{M7} - 420 mV). Formate is produced by anaerobes in the gut from mixed acid type fermentation reactions, so it is available to an organism like *C. jejuni*, which has the enzymatic capacity for formate dehydrogenation [21].

Weerakoon *et al.* [61] highlighted the importance of FDH in *C. jejuni* by creating a mutant with the chloramphenicol resistance gene inserted into a gene which codes for the large FDH subunit (*fdhA*, illustrated in Figure 7.3). The mutant had a 30-fold decrease in ability to respire using formate, however the growth was unaltered. The inactivation of FDH resulted in a statistically

significant decrease in the ability of the *C. jejuni fdhA* inactivated mutant to colonise the chicken caecum when compared to the wild-type. The natural flora of the chicken caecum is dominated by anaerobic fermentative organisms which provide an abundant source of both hydrogen and formate. Therefore, a mutant unable to oxidise either of these substrates in its ecological niche is most likely a costly disadvantage and ultimately may result in other micro-organisms out competing this *C. jejuni* FDH mutant strain for colonisation [61]. This investigation highlighted the importance of a fully-functioning FDH for *C. jejuni*.

In addition to its important role in colonisation, *C. jejuni* FDH is likely to be a tungstoenzyme. Smart *et al.* [64] engineered *C. jejuni* mutants with an inactive tungstate transporter (*tupA*). The *tupA* inactivated mutant had reduced FDH activity by approximately 50%, and when the molybate transporter (*modA* – which can also transport tungstate into the cell) was inactivated in addition to *tupA*, there was no FDH activity. This activity was restored by genetic complementation and the FDH activity in the wild-type could be enhanced by the addition of 1 mM sodium tungstate to the media. The medium used in this study was Müller–Hinton plus serine broth, which contained approximately 5 nM tungsten and 300 nM molybdenum. The amount of selenium (for the FDH selenoprotein) or formate in the growth medium was not reported. The results from the investigation suggested that *C. jejuni* possesses a specific, ultra-high affinity tungstate transporter that supplies tungsten for incorporation into FDH [64].

The genome sequence of *C. jejuni* NCTC 11168, reveals an operon encoding FDH subunits. FDH in *C. jejuni* is comprised of four subunits coded for by; *fdhA, fdhB, fdhC* and *fdhD* (Cj1511c-08c, Figure 7.3). The most notable of these is subunit *fdhA*; the largest subunit which encodes a 104 kDa selenocysteine containing tungstoprotein equivalent to the *E. coli* 110 kDa FdnG (α) subunit. Then, *fdhB* (24 kDa) and *fdhC* (35 kDa) are annotated as iron sulfur and cytochrome-B-containing, respectively. In addition to these three subunits, *C. jejuni* encode an FdhD protein (29 kDa) predicted to be required for activity of the FDH enzyme complex [36, 61]. In *E. coli*, the *fdhD* is required for the formation of FDH, but does not control the synthesis of FDH [65, 66].

Cj1514c was shown to be specifically required for the activity of FDH and was designated *fdhM*. [67]. Cj1513c is annotated as putative periplasmic protein and is a small gene of unknown function predicted to be TAT-transported (twin arginine translocation) out of the cell. Cj1513c is translationally coupled to the downstream *fdhA* gene and seems to be specific to the ε-proteobacteria. One intriguing possibility is that this protein is somehow related to the likely use of a tungsten-pterin cofactor in the *C. jejuni* formate dehydrogenase [67].

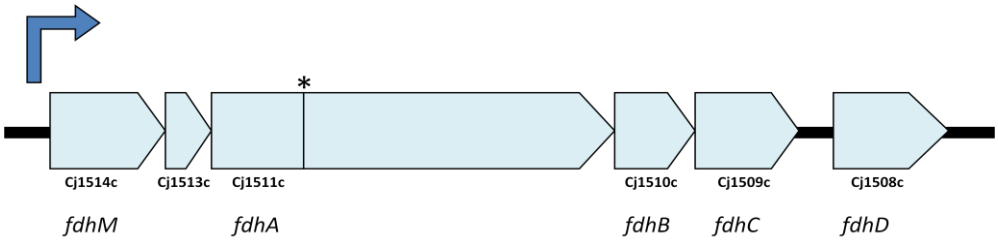


Figure 7.3 The FDH operon in *C. jejuni* 11168; *fdhA-D* and *fdhM*
*the position of the selenocysteine (181 aa), - the promotor

The *C. jejuni* FDH is a selenoprotein. This is known because of the sequence homology of the *fdhA* (Cj1511c, YP_002344890.1, Gene ID: 905795, Protein entry: EC 1.2.1.2) subunit to the selenocysteine-containing FDH proteins from *E. coli* and *Wolinella succinogenes*. Selenoproteins are relatively unusual for a number of reasons; they contain the amino acid selenocysteine, which is not commonly available intracellulary in bacteria, and selenocysteine is coded for by the codon UGA which would normally be read as a 'stop' codon [68].

7.3.1.1 Selenoproteins

During the early 1970s, one mammalian and two bacterial enzymes were identified as containing selenium as an integral constituent [68, 69]. This was shortly followed by the identification of the 21st amino acid: selenocysteine. Selenocysteine has the same structure as cysteine except that the sulphur atom is replaced by selenium. Selenocysteine has both a lower pKa and a higher reduction potential than cysteine.

Interestingly in selenoproteins, the in-frame UGA codon (which is normally a stop codon) is read as a selenocysteine [68-70]. This phenomenon takes place due to the role of a specialised element present in the mRNA, called SECIS

(SElenoCysteine Insertion Sequence), which is located immediately 3' of the UGA. This ensures the incorporation of the selenocysteine into the polypeptide chain. The SECIS element is defined by characteristic nucleotide sequences and secondary structure base-pairing patterns. Selenocysteine formation and incorporation have been extensively reviewed, and much is understood about the SECIS element, and the genes which interact with it, e.g. *selA* and *selB* [70-72].

Selenocysteine insertion is a unique example of co-translational insertion of a non-standard amino-acid. The dual nature of the UGA codon (a stop and a selenocysteine) raises an important question of the ability of the cell to distinguish between these two functions. The position of the SECIS element is evidence that a UGA is a selenocysteine codon, however, mis-annotations of sequenced genomes can, and do, occur. When cells are grown in the absence of selenium, translation of selenoproteins terminates at the UGA codon, resulting in a truncated, nonfunctional enzyme [73]. It can be hypothesised that a truncated protein that is produced by termination at the UGA codon could act as a dominant negative or have a regulatory function in the cell [74].

It can be argued that UGA is by far the most fascinating codon within the genetic code as it likely has served more functions than any other code in evolution. For example, an examination of current genetic language shows that UGA functions as a termination codon [75]; a selenocysteine codon [69]; a cysteine codon in *Euplotes octocarinatus* [76]; a tryptophan codon in mitochondria [77], *Mycoplasma* and *Sprioplasma* [77, 78]; an inefficiently read tryptophan codon in *Bacillus subtilis* [79]; and an inefficiently read codon in *E. coli* that is presumably decoded by tryptophan tRNA [80]. Another theory, discussed by Bock *et al.* [68], describes the original function of UGA to be for selenocysteine; a function which has mostly been lost during evolution, when oxygen levels increased in the earth's atmosphere. Selenocysteine is often found in micro-organisms from anaerobic or well-protected chemical environments, but the fact that it remains in both eukaryotes and prokaryotes indicates that the function developed before the lineages split.

There are no selenoproteins in yeast or higher plants [74]. In prokaryotes and archaebacteria, selenoproteins have been identified in only a limited number of species [81]. However, the presence of these unusual proteins across different prokaryotic species is noteworthy, and although the insertion of a selenocysteine seems like a labour-intensive procedure, the benefits of containing a selenoprotein have been conserved across bacterial evolution. Despite the functional diversity and scattered distribution of selenoproteins, many features of selenocysteine incorporation mechanism are conserved across species [74]. Notably, there is almost no overlap between the prokaryotic and eukaryotic selenoproteomes [74].

7.4 Concluding Summary

This study will investigate, using phenotypic investigations, the necessity for a fully-functioning *C. jejuni* FDH enzyme; an enzyme which is most likely a selenoprotein and has been reported to be necessary for *C. jejuni* colonisation in the chickens. Using genetic manipulation, *C. jejuni* strains will have genes which are required for FDH inactivated by the insertion of an antibiotic cassette. The strains will be tested phenotypically using techniques such as proteomics, transcriptomics and metabolomics. In addition to inactivation of certain genes, the strains will also be manipulated to express additional genes. These genes will be homologously recombined into the *C. jejuni* genome within a known pseudogene region.

This investigation will provide insight into a specific area of *C. jejuni* metabolism and allow further understanding of the pathogenesis of this organism. In addition, comprehension of the role of selenium and selenoproteins in *C. jejuni* will be investigated, and the findings regarding this interesting biosynthetic pathway may be applicable to other pathogenic and non-pathogenic micro-organisms.

Chapter Seven – References

1. Sebald M, Veron M: **Base DNA Content and Classification of Vibrios**. *Annales d'Institute Pasteur (Paris)* 1963, **105**:897-910.
2. Nachamkin I, Engberg J, Aarestrup F: **Diagnosis and Antimicrobial Suscepability of Campylobacter Species** In: *Campylobacter.* Edited by Blaser M, Nachamkin I, 2 edn. Washington: ASM Press 2000: 45.
3. DEFRA: **Zoonoses Report**. 2008, **PB13386**.
4. EFSA: **Scientific Opinion on Quantification of the risk posed by broiler meat to human campylobacteriosis in the EU**. *EFSA Journal* 2010, **8**(1).
5. Guerry P: **Campylobacter flagella: not just for motility**. *Trends in Microbiology* 2007, **15**(10):456-461.
6. Gundogdu O, Bentley SD, Holden MT, Parkhill J, Dorrell N, Wren BW: **Re-annotation and re-analysis of the *Campylobacter jejuni* NCTC11168 genome sequence**. *BMC Genomics* 2007, **8**:162.
7. Janssen R, Krogfelt KA, Cawthraw SA, van Pelt W, Wagenaar JA, Owen RJ: **Host-pathogen interactions in Campylobacter infections: the host perspective**. *Clinical Microbiology Reviews* 2008, **21**(3):505-518.
8. Young KT, Davis LM, Dirita VJ: ***Campylobacter jejuni*: molecular biology and pathogenesis**. *Nature Reviews Microbiology* 2007, **5**(9):665-679.
9. Redman N: **Food safety: a reference handbook** 2nd edn. Santa-Babara: ABC-CLIO; 2007.
10. FDA: **Bad Bug Book: Introduction Foodborne Pathogenic Microorganisms and Natural Toxins Handbook**. In.: U.S. Food and Drug Administration 2009.
11. Kothary M, Babu U: **Infective dose of foodborne pathogens in volunteers: A review**. *Journal of Food Safety* 2001, **21**(1):49-73.
12. Black RE, Levine MM, Clements ML, Hughes TP, Blaser MJ: **Experimental Campylobacter jejuni infection in humans**. *Journal of Infectious Disease* 1988, **157**(3):472-479.
13. CCDR: **Canadian Integrated Surveillance Report:Salmonella, Campylobacter, verotoxigenic *E. coli* and Shigella, from 2000 to**

2004. In., vol. Volume 35-S3: Human Campylobacter Cases Canada Communicable Disease Report (CCDR); 2009.

14. Lillehoj EP, Yun CH, Lillehoj HS: **Vaccines against the avian enteropathogens Eimeria, Cryptosporidium and Salmonella**. *Animal Health Research Reviews* 2000, **1**(1):47-65.
15. Tauxe RV: **Epidemiology of *Campylobacter jejuni* infections in the United States and other industrialized nations**. In: *Campylobacter jejuni: Current Status and Future Trends.* Edited by Nachamkin I, Blaser, M.J., Tompkins, L.S. . Washington, DC: American Society for Microbiology; 1992.
16. Nachamkin I, Allos BM, Ho T: **Campylobacter species and Guillain-Barre syndrome**. *Clinical Microbiology Reviews* 1998, **11**(3):555-567.
17. Hannu T, Mattila L, Rautelin H, Pelkonen P, Lahdenne P, Siitonen A, Leirisalo-Repo M: **Campylobacter-triggered reactive arthritis: a population-based study**. *Rheumatology (Oxford)* 2002, **41**(3):312-318.
18. Matsuda M, Moore JE: **Urease-positive thermophilic Campylobacter species**. *Appl Environ Microbiol* 2004, **70**(8):4415-4418.
19. Solomon E, Hoover D: ***Campylobacter jejuni:* A bacterial paradox**. *Journal of Food Safety* 1999, **19**(2):121-136.
20. Kelly D: **Metabolism, Electron Transport and Bioenergetics of *Campylobacter jejuni*: Implications for Understanding Life in the Gut and Survival in the Environment**. In: *Campylobacter: Molecular and Cellular Biology.* Edited by Konkel M, Ketley J. Wymondham: Horizon Bioscience; 2005.
21. Kelly D: **Complexity and Versatility in the Physiology and Metabolism of *Campylobacter jejuni*** In: *Campylobacter* Edited by Nachmkin I, Szymanski C, Blaser M, 3 edn. Washington DC: ASM Press; 2008.
22. Li X, Swaggerty CL, Kogut MH, Chiang HI, Wang Y, Genovese KJ, He H, Zhou H: **Gene expression profiling of the local cecal response of genetic chicken lines that differ in their susceptibility to *Campylobacter jejuni* colonization**. *PLoS One* 2010, **5**(7):e11827.
23. Gaynor EC, Wells DH, MacKichan JK, Falkow S: **The *Campylobacter jejuni* stringent response controls specific stress survival and**

virulence-associated phenotypes. *Molecular Microbiology* 2005, **56**(1):8-27.

24. van Putten JP, van Alphen LB, Wosten MM, de Zoete MR: **Molecular mechanisms of campylobacter infection**. *Current Topics in Microbiology and Immunology* 2009, **337**:197-229.
25. Chaveerach P, ter Huurne AA, Lipman LJ, van Knapen F: **Survival and resuscitation of ten strains of *Campylobacter jejuni* and *Campylobacter coli* under acid conditions**. *Applied and Environmental Microbiology* 2003, **69**(1):711-714.
26. Konturek JW, Thor P, Maczka M, Stoll R, Domschke W, Konturek SJ: **Role of cholecystokinin in the control of gastric emptying and secretory response to a fatty meal in normal subjects and duodenal ulcer patients**. *Scand J Gastroenterol* 1994, **29**(7):583-590.
27. Kelana LC, Griffiths MW: **Use of an autobioluminescent *Campylobacter jejuni* to monitor cell survival as a function of temperature, pH, and sodium chloride**. *Journal of Food Protection* 2003, **66**(11):2032-2037.
28. Murphy C, Carroll C, Jordan KN: **Induction of an adaptive tolerance response in the foodborne pathogen, *Campylobacter jejuni***. *FEMS Microbiol Letters* 2003, **223**(1):89-93.
29. Reid AN, Pandey R, Palyada K, Whitworth L, Doukhanine E, Stintzi A: **Identification of *Campylobacter jejuni* genes contributing to acid adaptation by transcriptional profiling and genome-wide mutagenesis**. *Applied and Environmental Microbiology* 2008, **74**(5):1598-1612.
30. Reid AN, Pandey R, Palyada K, Naikare H, Stintzi A: **Identification of *Campylobacter jejuni* genes involved in the response to acidic pH and stomach transit**. *Applied and Environmental Microbiology* 2008, **74**(5):1583-1597.
31. Richardson PT, Park SF: **Integration of heterologous plasmid DNA into multiple sites on the genome of Campylobacter coli following natural transformation**. *J Bacteriol* 1997, **179**(5):1809-1812.
32. Stewart GJ, Carlson CA: **The biology of natural transformation**. *Annu Rev Microbiol* 1986, **40**:211-235.

33. Wang Y, Taylor DE: **Natural transformation in Campylobacter species**. *J Bacteriol* 1990, **172**(2):949-955.
34. Taylor DE, Courvalin P: **Mechanisms of antibiotic resistance in Campylobacter species**. *Antimicrobial Agents and Chemotherapy* 1988, **32**(8):1107-1112.
35. Smith JL, Fratamico PM: **Fluoroquinolone resistance in campylobacter**. *Journal of Food Protection* 2010, **73**(6):1141-1152.
36. Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S *et al*: **The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences**. *Nature* 2000, **403**(6770):665-668.
37. Friis C, Wassenaar TM, Javed MA, Snipen L, Lagesen K, Hallin PF, Newell DG, Toszeghy M, Ridley A, Manning G *et al*: **Genomic characterization of *Campylobacter jejuni* strain M1**. *PLoS One* 2010, **5**(8):e12253.
38. Raskin DM, Seshadri R, Pukatzki SU, Mekalanos JJ: **Bacterial genomics and pathogen evolution**. *Cell* 2006, **124**(4):703-714.
39. Price EP, Thiruvenkataswamy V, Mickan L, Unicomb L, Rios RE, Huygens F, Giffard PM: **Genotyping of *Campylobacter jejuni* using seven single-nucleotide polymorphisms in combination with flaA short variable region sequencing**. *Journal of Medical Microbiology* 2006, **55**(Pt 8):1061-1070.
40. Parker CT, Miller WG, Horn ST, Lastovica AJ: **Common genomic features of *Campylobacter jejuni* subsp. *doylei* strains distinguish them from *C. jejuni* subsp. *jejuni***. *BMC Microbiology* 2007, **7**:50.
41. Gaskin DJ, Reuter M, Shearer N, Mulholland F, Pearson BM, van Vliet AH: **Genomics of thermophilic campylobacter species**. *Genome dynamics* 2009, **6**:91-109.
42. Bereswill S, Kist M: **Molecular microbiology and pathogenesis of Helicobacter and Campylobacter updated: a meeting report of the 11th conference on Campylobacter, Helicobacter and related organisms**. *Molecular Microbiology* 2002, **45**(1):255-262.
43. Carulli JP, Artinger M, Swain PM, Root CD, Chee L, Tulig C, Guerin J, Osborne M, Stein G, Lian J *et al*: **High throughput analysis of**

differential gene expression. *Journal of Cellular Biochemistry* 1998, **30-31**:286-296.

44. Atack JM, Kelly DJ: **Oxidative stress in *Campylobacter jejuni*: responses, resistance and regulation**. *Future Microbiology* 2009, **4**(6):677-690.
45. Park S: ***Campylobacter jejuni* Stress Responses During Survival in the Food Chain and Colonisation**. In: *Camplyobacter Molecular and Cellular Biology* Edited by Ketley J, Konkel M. Wymondham Horizon Bioscience 2005.
46. Konkel ME, Kim BJ, Klena JD, Young CR, Ziprin R: **Characterization of the thermal stress response of *Campylobacter jejuni***. *Infection and Immunity* 1998, **66**(8):3666-3672.
47. Stintzi A: **Gene expression profile of *Campylobacter jejuni* in response to growth temperature variation**. *Journal of Bacteriology* 2003, **185**(6):2009-2016.
48. Murphy C, Carroll C, Jordan KN: **Environmental survival mechanisms of the foodborne pathogen *Campylobacter jejuni***. *Journal of Applied Microbiology* 2006, **100**(4):623-632.
49. Giuliodori AM, Gualerzi CO, Soto S, Vila J, Tavio MM: **Review on bacterial stress topics**. *Ann N Y Acad Sci* 2007, **1113**:95-104.
50. Hazeleger WC, Wouters JA, Rombouts FM, Abee T: **Physiological activity of *Campylobacter jejuni* far below the minimal growth temperature**. *Applied and Environmental Microbiology* 1998, **64**(10):3917-3922.
51. Chaikam V, Karlson DT: **Comparison of structure, function and regulation of plant cold shock domain proteins to bacterial and animal cold shock domain proteins**. *Biochemistry and Molecular Biology Reports* 2010, **43**(1):1-8.
52. Moen B, Oust A, Langsrud O, Dorrell N, Marsden GL, Hinds J, Kohler A, Wren BW, Rudi K: **Explorative multifactor approach for investigating global survival mechanisms of *Campylobacter jejuni* under environmental conditions**. *Applied and Environmental Microbiology* 2005, **71**(4):2086-2094.

53. Ma Y, Hanning I, Slavik M: **Stress-Induced Adaptive Tolerance Response and Virulence Gene Expression in *Campylobacter jejuni***. *Journal of Food Safety* 2009, **29**(1):126-143.

54. Haddad N, Marce C, Magras C, Cappelier JM: **An overview of methods used to clarify pathogenesis mechanisms of *Campylobacter jejuni***. *Journal of Food Protection* 2010, **73**(4):786-802.

55. Flint HJ, O'Toole PW, Walker AW: **Special issue: The Human Intestinal Microbiota**. *Microbiology* 2010, **156**(Pt 11):3203-3204.

56. Thomas MT, Shepherd M, Poole RK, van Vliet AH, Kelly DJ, Pearson BM: **Two respiratory enzyme systems in *Campylobacter jejuni* NCTC 11168 contribute to growth on L-lactate**. *Environmental Microbiology* 2011, **13**(1):48-61.

57. Velayudhan J, Kelly DJ: **Analysis of gluconeogenic and anaplerotic enzymes in *Campylobacter jejuni:* an essential role for phosphoenolpyruvate carboxykinase**. *Microbiology* 2002, **148**(3):685-694.

58. Westfall HN, Rollins DM, Weiss E: **Substrate utilization by *Campylobacter jejuni* and *Campylobacter coli***. *Applied and Environmental Microbiology* 1986, **52**(4):700-705.

59. Leach S, Harvey P, Wali R: **Changes with growth rate in the membrane lipid composition of and amino acid utilization by continuous cultures of *Campylobacter jejuni***. *Journal of Applied Microbiology* 1997, **82**(5):631-640.

60. Myers JD, Kelly DJ: **A sulphite respiration system in the chemoheterotrophic human pathogen *Campylobacter jejuni***. *Microbiology* 2005, **151**(1):233-242.

61. Weerakoon DR, Borden NJ, Goodson CM, Grimes J, Olson JW: **The role of respiratory donor enzymes in *Campylobacter jejuni* host colonization and physiology**. *Microbial Pathogenesis* 2009, **47**(1):8-15.

62. Hoffman PS, Goodman TG: **Respiratory physiology and energy conservation efficiency of *Campylobacter jejuni***. *Journal of Bacteriology* 1982, **150**(1):319-326.

63. Popov VO, Lamzin VS: **NAD(+)-dependent formate dehydrogenase**. *Biochemistry Journal* 1994, **301 (Pt 3)**:625-643.

64. Smart JP, Cliff MJ, Kelly DJ: **A role for tungsten in the biology of *Campylobacter jejuni*: tungstate stimulates formate dehydrogenase activity and is transported via an ultra-high affinity ABC system distinct from the molybdate transporter**. *Molecular Microbiology* 2009, **74**(3):742-757.

65. Schlindwein C, Giordano G, Santini CL, Mandrand MA: **Identification and expression of the *Escherichia coli fdhD* and *fdhE* genes, which are involved in the formation of respiratory formate dehydrogenase**. *Journal of Bacteriology* 1990, **172**(10):6112-6121.

66. Stewart V, Lin JT, Berg BL: **Genetic evidence that genes *fdhD* and *fdhE* do not control synthesis of formate dehydrogenase-N in *Escherichia coli* K-12**. *Journal of Bacteriology* 1991, **173**(14):4417-4423.

67. Hitchcock A, Hall SJ, Myers JD, Mulholland F, Jones MA, Kelly DJ: **Roles of the twin-arginine translocase and associated chaperones in the biogenesis of the electron transport chains of the human pathogen *Campylobacter jejuni***. *Microbiology* 2010, **156**(10):2994-3010.

68. Bock A, Forchhammer K, Heider J, Leinfelder W, Sawers G, Veprek B, Zinoni F: **Selenocysteine: the 21st amino acid**. *Molecular Microbiology* 1991, **5**(3):515-520.

69. Bock A, Forchhammer K, Heider J, Baron C: **Selenoprotein synthesis: an expansion of the genetic code**. *Trends in Biochemical Sciences* 1991, **16**(12):463-467.

70. Hatfield DL, Gladyshev VN: **How selenium has altered our understanding of the genetic code**. *Molecular and Cellular Biology* 2002, **22**(11):3565-3576.

71. Walczak R, Hubert N, Carbon P, Krol A: **Solution structure of SECIS, the mRNA element required for eukaryotic selenocysteine insertion--interaction studies with the SECIS-binding protein SBP**. *Biomedical and Environmental Sciences* 1997, **10**(2-3):177-181.

72. Driscoll DM, Copeland PR: **Mechanism and regulation of selenoprotein synthesis**. *The Annual Review of Nutrition* 2003, **23**:17-40.

73. Arner ES: **Recombinant expression of mammalian selenocysteine-containing thioredoxin reductase and other selenoproteins in *Escherichia coli***. *Methods in Enzymology* 2002, **347**:226-235.

74. Driscoll DM, Chavatte L: **Finding needles in a haystack. In silico identification of eukaryotic selenoprotein genes**. *EMBO Rep* 2004, **5**(2):140-141.

75. Nirenberg M, Caskey T, Marshall R, Brimacombe R, Kellogg D, Doctor B, Hatfield D, Levin J, Rottman F, Pestka S *et al*: **The RNA code and protein synthesis**. *Cold Spring Harbor Symposia on Quantitative Biology* 1966, **31**:11-24.

76. Meyer F, Schmidt HJ, Plumper E, Hasilik A, Mersmann G, Meyer HE, Engstrom A, Heckmann K: **UGA is translated as cysteine in pheromone 3 of Euplotes octocarinatus**. *PNAS* 1991, **88**(9):3758-3761.

77. Osawa S, Jukes TH, Watanabe K, Muto A: **Recent evidence for evolution of the genetic code**. *Microbiology Reviews* 1992, **56**(1):229-264.

78. Watanabe K, Osawa S: **tRNA sequences and variations in the genetic code**. In: *tRNA: structure, biosynthesis, and function.* Edited by Söll D, RajBhandary UL. Washington, D.C.: American Society for Microbiology; 1995.

79. Lovett PS, Ambulos NP, Jr., Mulbry W, Noguchi N, Rogers EJ: **UGA can be decoded as tryptophan at low efficiency in *Bacillus subtilis***. *Journal of Bacteriology* 1991, **173**(5):1810-1812.

80. Weiner AM, Weber K: **A single UGA codon functions as a natural termination signal in the coliphage q beta coat protein cistron**. *Journal of Molecular Biology* 1973, **80**(4):837-855.

81. Bock A: **Selenium metabolism in bacteria** In: *Selenium.* Edited by Hatfield, L D. Hingham, MA, USA: Kluwer Academic Publishers; 2001: 7-22.

Chapter Eight

The genes Cj1500-01 have a crucial role in Formate Dehydrogenase Activity in *Campylobacter jejuni*

TABLE OF TABLES

TABLE OF FIGURES

Chapter Eight - The genes Cj1500-01 have a crucial role in Formate Dehydrogenase Activity in *Campylobacter jejuni*

8.1 Introduction

Pathogenesis by the gastro-intestinal pathogen *Campylobacter jejuni* strain NCTC 11168 remains poorly understood [1, 2]. The unique presence and regulation of expression of certain genes can be linked to *C. jejuni* survival and pathogenicity. Understanding these genomic responses to stressful environmental conditions, will allow further comprehension of how the micro-organism adapts, survives and infects a host.

The small gene Cj1501 (228 bp, 75 amino acids, Figure 8.1) was previously identified as up-regulated in response to acid-shock (Figure 8.2, Ida Porcelli, unpublished data). The microarray study described genes differentially regulated during acid-shock. Cj1501 was identified as up-regulated over different acidic pHs (3.6 and 5.0) and exposure times (10 and 30 minutes) (Figure 8.2).

Cj1501 is an uncharacterised protein currently annotated as a 'hypothetical protein'. The 75 amino acid chain has sequence homology to the SirA domain (Figure 8.3, http://pfam.sanger.ac.uk), and is therefore classified as a SirA-like protein. In other bacteria, SirA proteins function as response regulators in a two-component regulatory system, often in conjunction with the sensor kinase, BarA [3]. However, there is an absence of a BarA homologue in *C. jejuni* and the Cj1501 lacks a recognisable DNA binding domain, characteristic of response regulators.

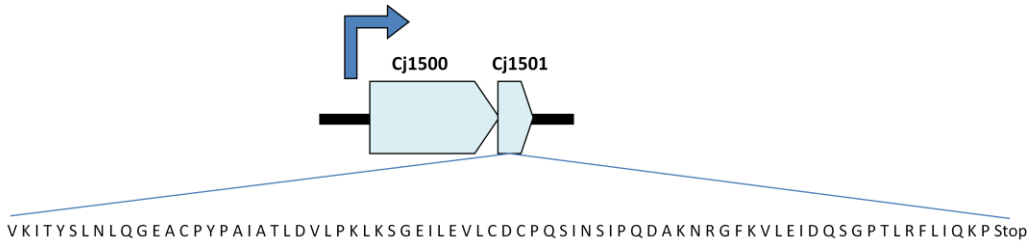


Figure 8.1 The amino acid sequence of Cj1501 and position of its promoter; upstream of Cj1500.
- the promotor

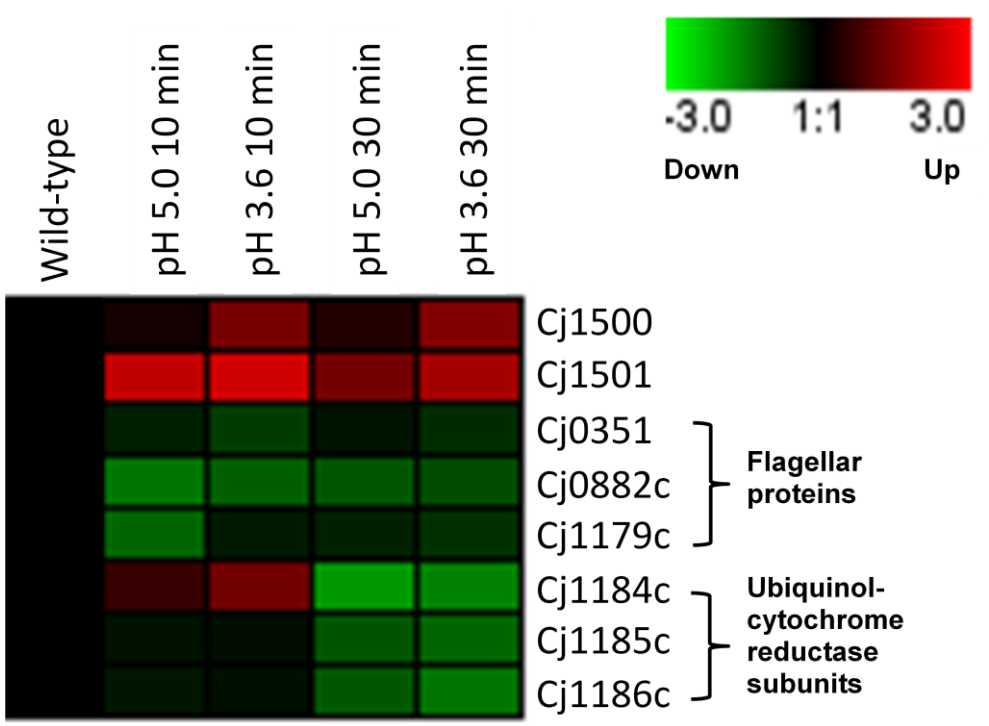


Figure 8.2 Heat map showing the up- and down-regulation of several *C. jejuni* genes
Following acid-shock at either pH 3.6 or 5.0 for either 10 or 30 minute exposures. The flagella accessory proteins and ubiquinol-cytochrome reductase subunits are shown for regulation comparison.

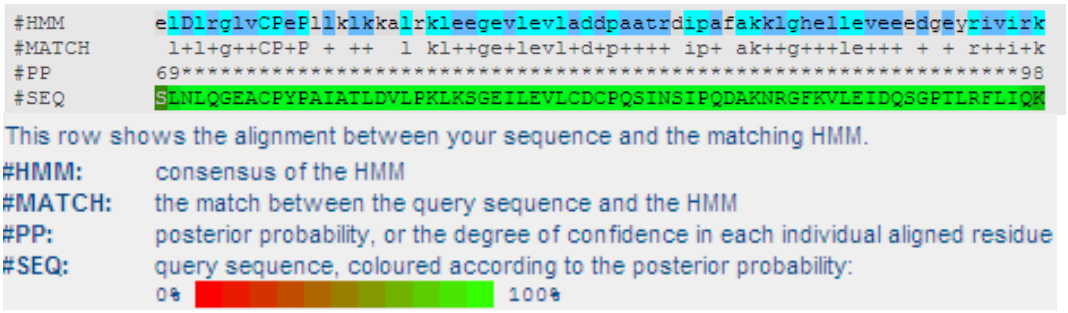


Figure 8.3 Shows the alignment of the Cj1501 with the predicted domain (SirA) generated by Pfam software.
HMM – Hidden Markov Model. The green colour indicates that this is a highly confident match.

Other phylogenetically diverse organisms, such as the β-Proteobacterium *Nitrosomonas europaea*, a Gram-negative obligate chemolithoautotroph, and the hyperthermophillic bacterium *Thermotoga petrophila*, contain homologues of the *C. jejuni* gene Cj1500 also adjacent to Cj1501 homologues. Because the two genes are found together over a variety of species, Cj1500 may be involved in the same regulatory, biosynthetic or metabolic pathway as the SirA-like protein, Cj1501. This distribution and the possible role of the Cj1500-1501 homologues is discussed in Chapter Ten.

Genomic analysis indicates that in *C. jejuni* strain 81116 Cj1500 and Cj1501 exist as a single fusion protein (C8J_1404) whereas in other *C. jejuni* strains, the genes are located sequentially and distinctly. Figure 8.4 shows this relationship between Cj1500 and Cj1501 orthologues in the Campylobacter species. The original sequencing data from the *C. jejuni* 81116 project [4] was re-examined to confirm the fusion of the two independent genes was genuine.

The role of Cj1501 was investigated by creating a suite of mutants, in two strains of *C. jejuni*, with the Cj1501 (or homologue, C8J_1404) and preceding gene, Cj1500, inactivated by insertion of an antibiotic resistance cassette. A list of mutants and complemented mutants created in this investigation are detailed in Table 8.1. The phenotype and proteome of the insertion mutants were studied under normal and acidic stress conditions.

Figure 8.4 A representation of the relationship between Cj1500 (red) and Cj1501 (beige) orthologues in the Campylobacter species.
Neighbourhood display - genes connected by black lines are in immediate neighbourhood on the genome (within 300 bp on the same strand). The image was generated by http://string-db.org/ and is based upon the location of genes spatially within the genome. The fused genes in *C. jejuni* 81116 are identified by the asterisk*. Confidence level 0.4 (medium). The phylogenetic tree is generated based on the whole genome similarities between *C. jejuni* strains. The *selA* and other selenocysteine synthetic genes in *C. jejuni* NCTC 11168 are not depicted on this diagram, as it shows the immediate neighbourhood only and the most highly conserved genes therein. The *selA* and *selB* in *C. jejuni* NCTC 11168 occur geographically over 100 genes away from Cj1500, which this diagram is centered upon.

Little investigation has been done into this small, but highly conserved, gene Cj1501. Small genes, such as this, can be crucial for gene regulation and bacterial survival, but are often over-looked. As so much is still unknown about *C. jejuni* pathogenicity, this particular gene, identified by the consistent up-regulation in acid stress conditions, is an important area for investigation. The highly conserved nature across *C. jejuni* evolution suggests it possibly holds a crucial role.

Name	Inactivated Gene(s)	Gene Annotation	*C. jejuni* Strain	Genetic Complementation
Δ1501	Cj1501	SirA-like	NCTC 11168	n.a.
Δ1500	Cj1500	putative inner membrane protein	NCTC 11168	n.a.
Δ1500-01	Cj1501 & Cj1500	putative inner membrane protein & SirA-like	NCTC 11168	n.a.
Δ1404	C8J_1404	putative inner membrane protein	81116	n.a.
Δ1378	Cj1378	selenium transferase, *selA*	NCTC 11168	n.a.
Δ1379	Cj1379	putative selenocysteine-specific elongation factor, *selB*	NCTC 11168	n.a.
Δ1378-79	Cj1378 & Cj1379	selenium transferase, *selA*, putative selenocysteine-specific elongation factor, *selB*	NCTC 11168	n.a.
Δ1511	Cj1511c	Putative formate dehydrogenase large subunit (Selenocysteine containing), *fdhA*	NCTC 11168	n.a.
Δ1501::1501	Cj1501	SirA-like	NCTC 11168	Cj1501 under *fdxA* promoter
Δ1500::1501	Cj1500	putative inner membrane protein	NCTC 11168	Cj1501 under *fdxA* promoter
Δ1500-01::1501	Cj1501 & Cj1500	putative inner membrane protein & SirA-like	NCTC 11168	Cj1501 under *fdxA* promoter
Δ*flaAB*	Cj1339c & Cj1338c	structural flagella proteins, *flaA* & *flaB*, used as a motility defective strain	NCTC 11168	n.a.
Δ*cetAB*	Cj1190c & Cj1189c	bipartate energy taxis response proteins, *cetA* & *cetB*, used as an aerotaxis defective strain	NCTC 11168	n.a.
WT::1501	n.a.	n.a.	NCTC 11168	Cj1501 under *fdxA* promoter

Table 8.1 The mutants and complements discussed in this chapter.
n.a. – not applicable

8.2 Objectives

- To understand more about the acid-resistance and metabolism of *C. jejuni* NCTC 11168, focusing on a single gene (Cj1501) and its biological relevance within the organism.

- To study the gene (Cj1501) across the Campylobacter species.

- To create a selection of *C. jejuni* insertion mutants to investigate the effect of inactive genes.

- To purify the protein Cj1501 to understand more about its function.

8.3 Results

8.3.1 Cj1501, Cj1500, Cj1500-01 and C8J_1404 insertion inactivation *C. jejuni* mutants have no phenotypic differences in growth, motility, aerotaxis, auto-agglutination or resistance to acid-shock.

Growth

The growth of *C. jejuni* was monitored by measuring the optical density of the culture at OD_{600nm}. The results were then plotted on a log scale graph. *C. jejuni* typically grows exponentially for several hours and the final OD_{600nm} is dependent on the strain and culture system used. For example during this investigation, *C. jejuni* 81116 can naturally reach an optical density of 0.8, whereas *C. jejuni* NCTC 11168 will usually reach approximately 0.4. When mutants Δ1501, Δ1500, Δ1500-01 and Δ1404 were grown in Brucella broth in micro-aerobic conditions, at 37°C, with constant agitation at 200 rpm, there was no difference between each mutant and the wild-type strain it was derived from (Figure 8.5). All the insertion mutants exhibited a typical *C. jejuni* growth pattern. This suggested that an insertion mutation in Cj1500, Cj1501 or C8J_1404 does not affect the growth of *C. jejuni*. Therefore, the genes (Cj1500, Cj1501 and C8J_1404) are likely to not be involved in growth related functions, such as cell division and replication.

Motility

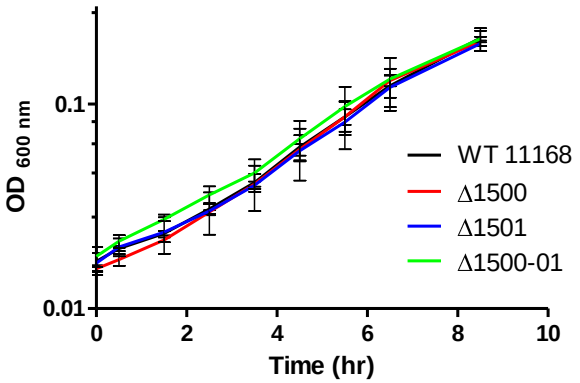
The motility of *C. jejuni* was measured when a 5 µl drop of liquid culture was placed on a soft agar plate in a circular form. If the *C. jejuni* are motile, they move across the surface of the agar in equal directions and the distance travelled is measured by recording the diameter of visible cells. This experiment tests the chemotatic ability of the *C. jejuni*, as well as the motility. When the motility of mutants Δ1501, Δ1500 & Δ1500-01 was compared with the wild-type strain (*C. jejuni* NCTC 11168), there was no significant difference in motility seen. This suggested that an insertion mutation in Cj1500 or Cj1501 does not affect the motility of *C. jejuni*, therefore, the genes are likely not involved in motility related functions, such as flagella assembly. Naturally, *C. jejuni* NCTC 11168 is a motile strain. The motility assay is fairly variable,

but to compensate, several biological and technical repeats were performed. As a negative control, a motility mutant (Δ*flaAB* – as referred to in Table 8.1) was used for comparison. The Δ*flaAB* bacteria did not spread across the plate, and remained in their droplet of approximate diameter 6 mm (Figure 8.6).

Aerotaxis

Naturally, *C. jejuni* will move away from areas of high oxygen concentration [5]. Although it is not fully understood, the aerotaxis assay is based on the assumption that oxygen diffuses into a tube of soft agar; the oxygen concentration at the top of the agar will be near atmospheric concentrations, whereas at the bottom of the tube, the oxygen concentration lower, as an oxygen gradient is formed throughout. The movement toward a favourable concentration of oxygen was expected. To measure the aerotaxis of *C. jejuni* the distance the bacteria move through soft agar away from the higher oxygen concentrations can be compared between mutants, wild-type and an aerotaxis defective mutant (Δ*cetAB*). This method showed that inactivation of the genes Cj1501 and C8J_1404 in *C. jejuni* strains NCTC 11168 and 81116 respectively, and a formate dehydrogenase defective mutant (Δ1511 derived from *C. jejuni* NCTC 11168) resulted in a slight decrease in aerotaxis. However, the phenotype seen in Δ1501 could not be restored by the re-expression of the gene (Δ1501::1501). This indicates that the phenotype was not due to the gene inactivation, but possibly an indirect effect of the mutagenesis. Interestingly, the phenotype was seen in both strains (Figure 8.7). The chemotaxis and aerotaxis of *C. jejuni* are areas of research which are not fully understood, and this result provides an interesting insight into this field. Although conclusions about the relative ability of the Δ1501 and Δ1404 to move through the agar cannot be determined based on this experiment alone, it would be an interesting area to continue to investigate, as chemotaxis and aerotaxis abilities may possibly be linked to *C. jejuni* pathogenesis. The two *C. jejuni* strains, 81116 and NCTC 11168, are able to move through the agar at different rates, this is why representative day three and seven are shown respectively to illustrate the phenotype. The colour change observed occurs by the action of various dehydrogenases in the *C. jejuni* cells. It is an irreversible reaction and represents where *C. jejuni* have been, not where they are currently respiring.

a



b

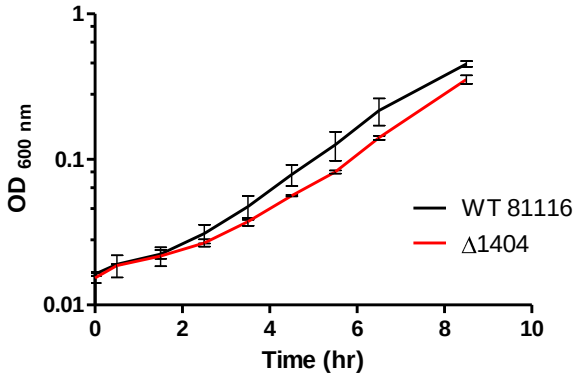


Figure 8.5 Post mutagenesis growth of *C. jejuni*
The growth of Δ1501, Δ1500, Δ1500-01 compared to the wild-type strain *C. jejuni* NCTC 11168 (a) and Δ1404 compared to the wild-type strain *C. jejuni* 81116 (b). The data is based on three biological repeats and the standard deviation is shown. The small differences between samples are not significant or relevant for this investigation.

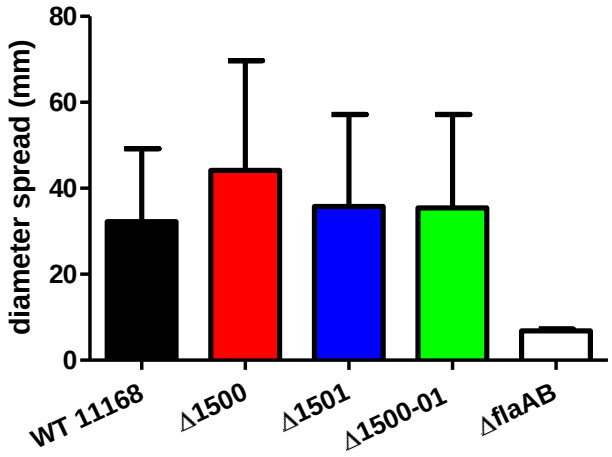


Figure 8.6 Post mutagenesis motility of *C. jejuni*
The motility of strains Δ1501, Δ1500, Δ1500-01 and the wild-type strain *C. jejuni* NCTC 11168 compared with a *C. jejuni* strain lacking motility (Δ*flaAB*). The motility of the strains is measured by the spread of the cells in mm on soft agar plates and the standard deviation is shown.

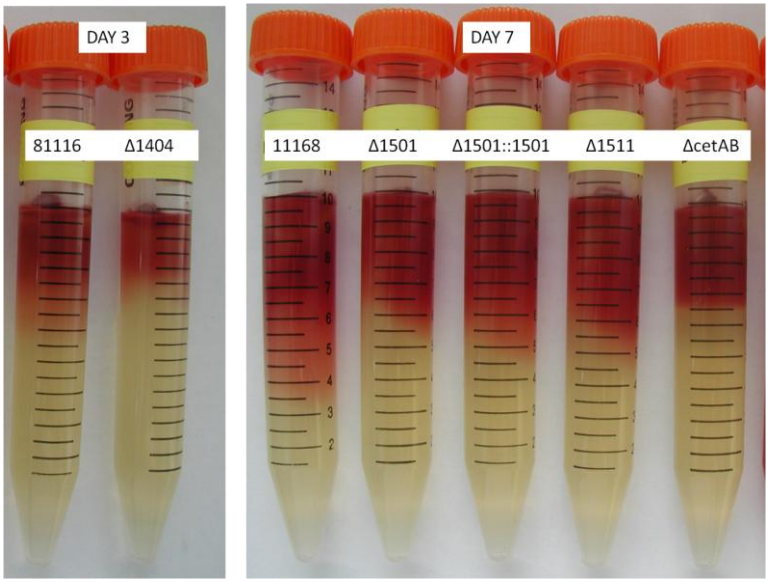


Figure 8.7 *C. jejuni* aerotaxis
The aerotaxis of *C. jejuni* 81116 wild-type and mutant Δ1404 after three days (left) and of *C. jejuni* NCTC 11168 wild-type, mutant Δ1501, restored mutant Δ1501::1501, formate dehydrogenase mutant Δ1511, and aerotaxis mutant Δ*cetAB*, after seven days (right). The colour change in the soft agar is due to the addition of 2,3,5-triphenyltetrazolium chloride (TTC), which becomes red (TPF,1,3,5-triphenylformazan) due to the activity of various cellular dehydrogenases which are active when the bacteria respire. The colour change is used to determine the distance the bacteria have travelled.

Auto-agglutination

The importance of auto-agglutination in virulence has been strongly implicated for other pathogenic bacteria, including *Yersinia enterocolitica*, enteropathogenic *Escherichia coli* and *Vibrio cholerae* [6-8]. Auto-agglutination is the tendency of bacteria to adhere to one another; as the bacteria clump together, these clumps gradually sink in the liquid medium. This effect can be monitored by OD_{600nm}. If a strain auto-agglutinates, the OD_{600nm} (at the top of the culture) will quickly fall as the bacteria sink. The flagella of *C. jejuni* are involved in auto-agglutination [9], but currently, the role of auto-agglutination in *C. jejuni* pathogenesis has not been determined [10]. The two strains of *C. jejuni* used in this investigation (81116 and NCTC 11168) show different auto-agglutination properties; 81116 does not auto-agglutinate, whereas NCTC 11168 does. The mutants developed in this investigation (Δ1501, Δ1500 & Δ1500-01 and Δ1404) have no variance in auto-agglutination ability when compared to the strain they were derived from (Figure 8.8). This suggests that an insertion mutation in Cj1500, Cj1501 or C8J_1404 does not affect the auto-agglutination of *C. jejuni*, therefore, the genes are likely not involved in auto-agglutination related functions, such as flagella assembly, because expression of flagella is correlated with auto-agglutination [10].

Resistance to Acid-Shock

Survival of *C. jejuni* at acid pH values is temperature dependent, but cellular inactivation is rapid at pH values less than 4.0, especially above refrigeration temperatures [11]. Initially, Cj1501 was identified as up-regulated during acid-shock (see Introduction). So, the ability of a mutant lacking the Cj1501 gene to survive acidic conditions was investigated. Cultures of *C. jejuni* NCTC 11168 and Δ1501 grown to both log- and stationary-phase were resuspended in acidic Brucella broth (pH 5.0, normal Brucella broth is pH 7.0). The cultures were incubated for 30 minutes in the micro-aerobic cabinet, at 37°C, with constant agitation at 200 rpm. Then, serial dilutions were plated onto Brucella plates. Unfortunately, no distinct, repeatable phenotype was observed during these investigations. It was concluded that inactivation mutation of Cj1501 had neither a positive nor a negative effect on the ability of *C. jejuni* to survive a low pH shock (Figure 8.9). These data in Figure 8.9a appear to show that at

pH 3.5 the wild-type is able to survive better than the Δ1501, however this result was not reproduceable. Previous work in the Campylobacter group at IFR suggested that survival of *C. jejuni* was drastically lost around pH 3.5. So testing *C. jejuni* endurance at pH 3.5 may have given contrasting results, as it is the cusp of *C. jejuni* survival. Figure 8.9b shows graphically the recovery of the *C. jejuni* cells after acid shock. At pH 3.5 the results varied greatly between biological and technical repeats; sometimes 100% recovery was seen, and sometimes complete death (0%), when compared with resuspension in pH 7.0 broth. Even though there was no distinct survival phenotype observed by an insertion mutant of Cj1501, the proteome of acid-shocked mutant Cj1501 was compared to that from an acid-shocked wild-type *C. jejuni* (discussed in section 8.3.3). It was hypothesised that if there were any explanations for the Cj1501 gene to be up-regulated upon acid-shock, they may have been visible on the proteome level.

In addition, the resistance to acid-shocking was tested using Δ1404 (derived from *C. jejuni* 81116) to determine if a definitive phenotype was present in this strain. Once again, the data was not reproducible, and the same variation of survival around pH 3.5 was observed (data not shown). Finally, to determine if the expression of Cj1501 in a Cj1501 insertion inactivated mutant (Δ1501::1501) could provide a similar phenotype to that of the wild-type, when compared with the Cj1501 inactivated mutant (Δ1501), the same experiment was performed. Unfortunately, no clear phenotype regarding the ability of the cells to withstand acid-shock could be attributed to the presence, over-expression, or absence of Cj1501 in *C. jejuni* NCTC 11168 using this experiment alone.

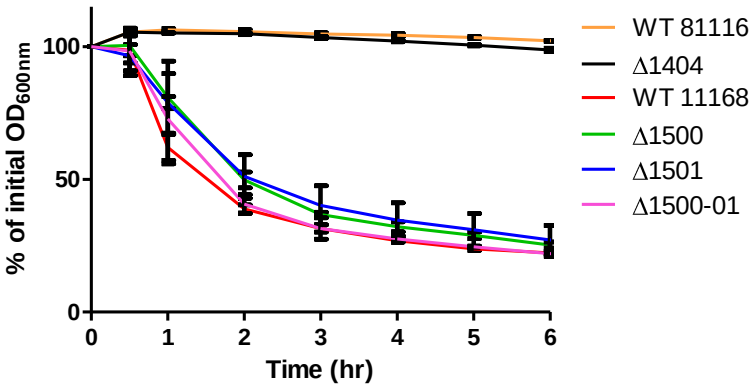


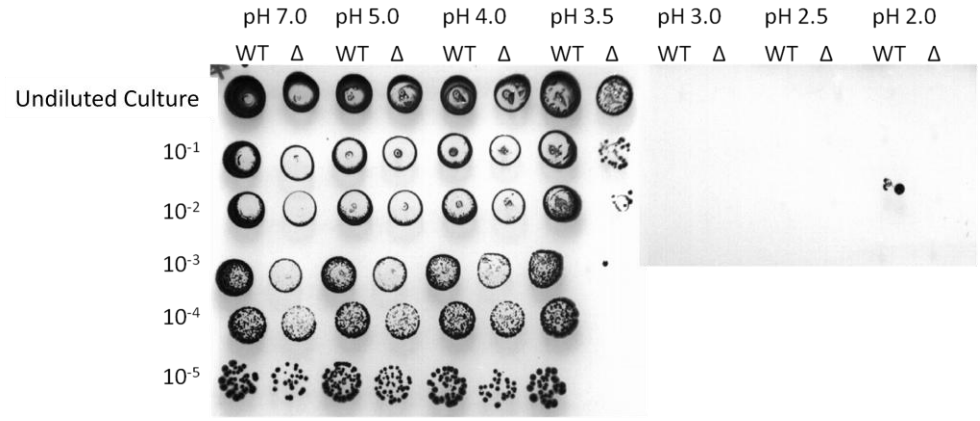
Figure 8.8 *C. jejuni* auto-agglutination
Shows the auto-agglutination of Δ1501, Δ1500 & Δ1500-01 compared with *C. jejuni* NCTC 11168 and Δ1404 compared with *C. jejuni* 81116. The 81116 and Δ1404 both did not auto-agglutinate under these conditions. The 11168 and Δ1501, Δ1500 & Δ1500-01 all have auto-agglutination properties. There was no difference between any mutant and its corresponding wild-type strain. The data is based on several biological repeats. The standard deviation is shown.

8.3.2 Intermediate Summary

There was no difference in growth, motility, aerotaxis, auto-agglutination and resistance to acid-shock between the mutants (Δ1501, Δ1500 & Δ1500-01 and Δ1404) and their corresponding wild-type strain. Therefore, it can be concluded that the genes Cj1501, Cj1500 and C8J_1404 are most likely not involved in functions such as cell growth and division, motility and flagella assembly, or acid resistance.

Any significant and repeatable differences observed between the mutant (Δ1501, Δ1500 & Δ1500-01 and Δ1404) and wild-type in the following experiments are definitely due to the mutation, and not an artifact of the insertion cassette. Sometimes, if an insertion mutation causes a phenotypic growth defect (or similar), the changes in proteome may be due to the growth (or similar) phenotype, and not the inactivation of the gene of interest. However, this was not the case in this investigation.

a



b

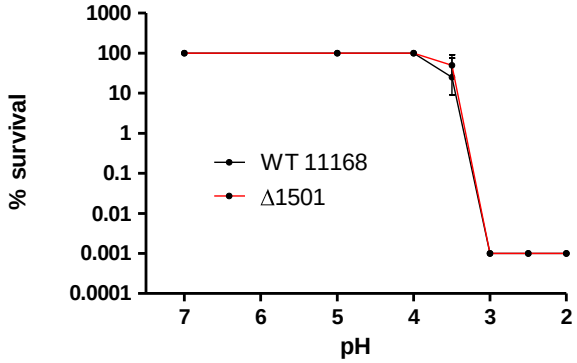


Figure 8.9 *C. jejuni* resistance to acid-shock
Representative data shown.
(a) Shows the growth recovery after acid-shock of the wild-type *C. jejuni* NCTC 11168 (WT) and Δ1501 (Δ) at several pHs. The ability to survive post acid-shock, measured by colony counts following serial dilutions, varied several-fold between repeat (biological and technical) experiments. These data here suggests that at pH 3.5 the wild-type was able to survive better than the Δ1501, however this result was not reproduceable.
(b) Using the colony count data, a graph shows the percentage survival of *C. jejuni* NCTC 11168 and Δ1501 after shocking cultures at acidic pHs compared with resuspension at pH 7.0. The *C. jejuni* cells (both NCTC 11168 and Δ1501) have 100% survival until pH 4.0, any pH below that, survival is drastically decreased. The cusp of *C. jejuni* survival is approximately pH 3.5. At this pH the variation between experiments was large; ranging from none (shown here at 0.001 on the scale) to full (100). The data for this graph is from cultures tested at stationary phase. The standard deviation is shown on the graph.

8.3.3 Proteomics revealed the Cj1501 inactivation mutant lacked the formate dehydrogenase protein subunits

With the exception of some outer-membrane proteins, the proteome of *C. jejuni* can be seen using a two-dimensional (2D) gel. The total cellular protein was extracted from *C. jejuni* (both wild-type NCTC 11168 and Δ1501), then separated based on molecular mass and isoelectrical focusing point. The proteins were stained with Sypro Ruby for imaging, and then images were overlaid with one another to compare and contrast the proteomes.

The wild-type *C. jejuni* NCTC 11168 proteome was compared to that of Δ1501, both under normal growth conditions, and after an acid-shock (re-suspension in pH 5.0 Brucella broth for 30 minutes under normal growth conditions). The acid-shock was chosen to be pH 5.0 because it was predicted that acid-response genes would be up-regulated at this pH, and therefore the resulting protein would be seen on the gel. If the campylobacter cells were greatly stressed by the pH shock, proteins seen on the gel may have been due to a general survival response, rather than specific to the pH environment. There were no repeatable and significant differences between the proteome of wild-type *C. jejuni* NCTC 11168 with and without acid-shock, or Δ1501 with and without acid-shock. This result suggests that any responses that *C. jejuni* has to survive acid-shock were not identified by this experiment, and were not influenced by the presence or absence of Cj1501. This result agreed with the data which suggested that the inactivation mutation of Cj1501 had neither a positive or negative effect on the ability of *C. jejuni* to survive a low pH shock (Figure 8.9).

When the proteome of Δ1501 was compared to the wild-type *C. jejuni* NCTC 11168 under normal growth conditions, a few proteins were identified as constantly absent in the Cj1501 inactivation insertion mutant (Δ1501) based on three biological repeat experiments (Figure 8.10). These proteins were removed from the wild-type gel, and digested with trypsin, and subjected to mass spectroscopy to determine the amino acid order of the protein. Comparison of the protein sequence to known databases allowed an accurate identification of each protein.

The constantly missing protein subunits from the Δ1501 mutant identified in this way were: Cj1511c; putative formate dehydrogenase large subunit (selenocysteine containing) *fdhA*, EC=1.2.1.2 and Cj1510c; putative formate dehydrogenase iron-sulfur subunit, *fdhB*, EC=1.2.1.2. Both proteins have been highlighted in Figure 8.10. Formate dehydrogenase (FDH) in *C. jejuni* is essential for optimum chicken caecum colonisation [12]. In addition, FDH is a predicted selenoprotein and is most likely a tungstoenzyme [13]. The lack of the FDH subunits was interesting because it was unknown that Cj1501 was related to the FDH protein of *C. jejuni*. The proteomic analysis indicated that when the gene Cj1501 was inactivated, the *C. jejuni* was unable to synthesise two of the FDH protein subunits. The rest of the proteome remains similar in the Δ1501 to the wild-type *C. jejuni*. It was hypothesised, based on this result, that Cj1501 had a role in FDH formation or stabilisation. Cj1501, at the start of this investigation, was not classified as having a role regarding FDH, or another protein, formation, so the result was interesting and required further analysis and investigation.

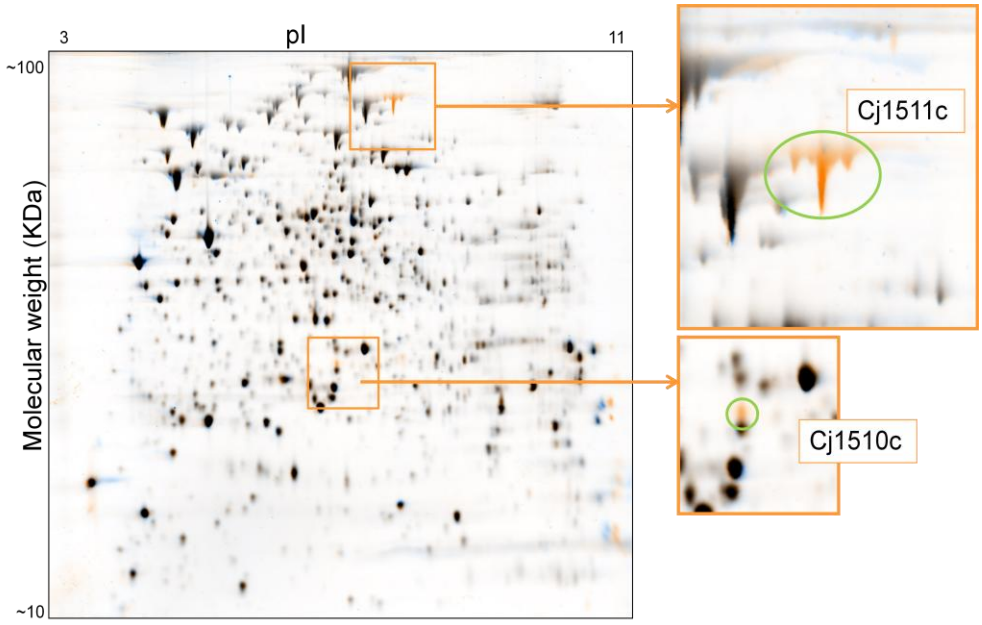


Figure 8.10 Proteomic analysis
Shows the overlay images of the proteome of wild-type *C. jejuni* NCTC 11168 (**orange**) and Δ1501 (**blue**). If the images match up exactly when overlaid, the proteins appear **black**. Clearly, proteins present in the wild-type, and absent in the mutant, can be identified (**orange**). The two proteins of interest have been highlighted and are shown to the right of the main image. Towards the right-hand side of the main overlay image, proteins which are not aligning exactly can be seen (replicate blue and orange spots). However, these proteins were not seen consistently during the biological repeats and the higher isoelectrical focusing points (pI) often have slight batch variations, which become apparent during the first dimension. A representative pair of gels is shown, although the result was reproducible over three biological repeats.

8.3.4 The Cj1501 insertion inactivation mutant lacked detectable formate dehydrogenase (FDH) activity, confirmed by benzyl viologen linked assay on living cells.

Following on from the proteomics result, which suggested Δ1501 lacked the FDH enzyme because two of the subunits were not visible after 2D proteomic seperation, a simple assay to determine FDH activity in living *C. jejuni* cells was performed. The reaction took place in an anaerobically sealed cuvette and was based upon the colour change of a coupled indicator, benzyl viologen [13]. The benzyl viologen changes from clear to purple on receipt of an electron during the dehydrogenation of a formate molecule:

$$\textbf{Formate} + NAD^+ + H^+ + 2e^- \leftrightarrow CO_2 + \textbf{NADH} + 2H^+ + 2e^-$$

The colour change was monitored spectrophotometrically at 578 nm in real time (one reading per second). The amount of colour change per minute, per mg protein, can be used to calculate a numerical activity for the FDH enzyme.

It was quickly established that in Δ1501 (derived from *C. jejuni* NCTC 11168), and the corresponding Δ1404 strain (derived from *C. jejuni* 81116), there was no detectable FDH activity, when compared to the respective wild-type strains (Figure 8.11). It is possible that if FDH was functioning in the cells, it may be below the level of detection for this coupled assay method. This result suggests that Cj1501 (or C8J_1404) is essential for FDH activity, and this result was in agreement with the 2D proteomic data, which showed the insertion mutant Δ1501 to lack two of the FDH subunits. The fact that both *C. jejuni* strains NCTC 11168 and 81116 with the corresponding homologues inactivated (Cj1501 and C8J_1404) lacked a detectable FDH activity, suggested that this phenotype was not strain-specific in *C. jejuni*.

Interestingly, this FDH-null phenotype was observed in Δ1501, Δ1500 and Δ1500-01 when compared with the wild-type (Figure 8.12). However, when Cj1500 was inactivated by insertion mutation, it is possible that the downstream Cj1501 may have been inactivated as a polar effect due to inhibition of transcription. To determine whether a disrupted Cj1500 was responsible for FDH inactivation, it was necessary to genetically complement

the mutants Δ1501, Δ1500 and Δ1500-01 with the gene Cj1501 (Δ1501::1501, Δ1500::1501, Δ1500-01::1501). The complemented mutants would provide information as to whether the FDH inactivity was due to an inactivated Cj1501, Cj1500, or both Cj1500-01. The mutants and their complements were tested for FDH activity (Figure 8.12). The complementation of Δ1501 with Cj1501 (Δ1501::1501) restored FDH activity to approximately 60% that of the wild-type strain. This suggested that the lack of FDH activity was due to the inactivation of Cj1501. In addition, complementation of Δ1500 and Δ1500-01 with Cj1501 (Δ1500::1501 and Δ1500-01::1501) did not restore FDH activity at all. This suggested that Cj1500 is also a crucial gene in the production of a functional FDH in *C. jejuni* NCTC 11168 and strengthens the bioinformatic link between Cj1500 and Cj1501. This was an interesting result because a function had never been assigned to Cj1500, and no links between Cj1500 and FDH activity have previously been established.

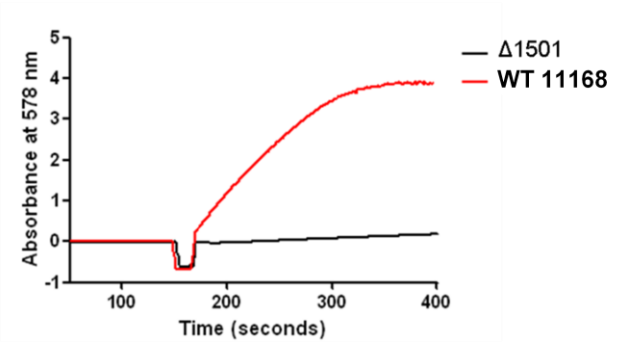


Figure 8.11 FDH activity (colour change)
The FDH activity, based on the benzyl viologen coupled-assay to visualise the reaction (one reading per second), of the wild-type (WT) *C. jejuni* (NCTC 11168), **red**, compared with a mutant Δ1501, **black**. The same phenomenon was seen with wild-type *C. jejuni* (81116) and the corresponding mutant Δ1404 (data not shown). The rate of reaction was calculated whilst the resulting curve was progressing in a linear fashion.

Further to genetic complementation, a method for chemically complementing the mutants (Δ1501, Δ1500 and Δ1500-01) was investigated. Because the FDH subunit missing during 2D proteomics of Δ1501 was selenocysteine-containing (*fdhA*), the addition of exogenous selenium to the growth media was tested to see if the FDH activity could be restored. In all mutant cases (Δ1501, Δ1500 and Δ1500-01) the FDH activity could be partially restored by the addition of exogenous 5 μM selenium (Figure 8.12). This result suggested

that perhaps the pair of sequential genes (Cj1500-01) could be involved with the incorporation of the selenium, or selenocysteine, into the FDH protein. Cj1500-01 may work as the transporter or apparatus to allow selenium incorporation into FDH. It has been shown in *E. coli* that when cells are grown in an absence of selenium, a short, non-functional selenoprotein is formed because translation was halted at the selenocysteine codon, UGA [14]. Therefore, it was hypothesised that perhaps a lack of intracellular selenium in *C. jejuni* mutant strains Δ1501, Δ1500 and Δ1500-01 was responsible for the lack of FDH activity. The result suggests that perhaps a large concentration of selenium activates alternative transporters or apparatus in *C. jejuni* variants Δ1501, Δ1500 and Δ1500-01, which acquire the selenium from the growth medium, synthesise the selenocysteine, and incorporate it into the FDH protein.

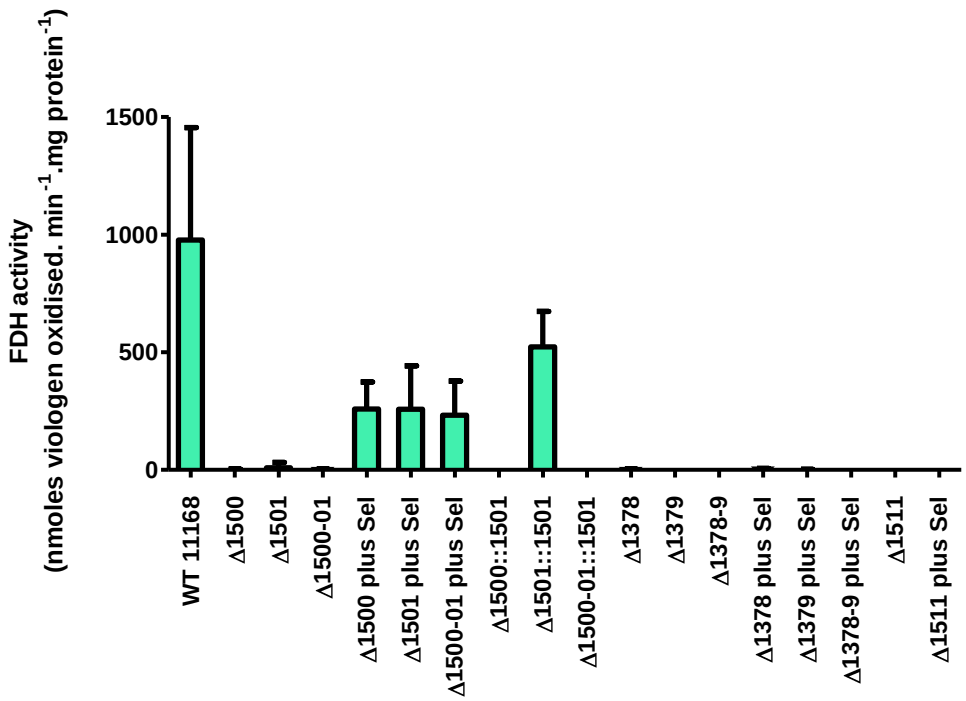


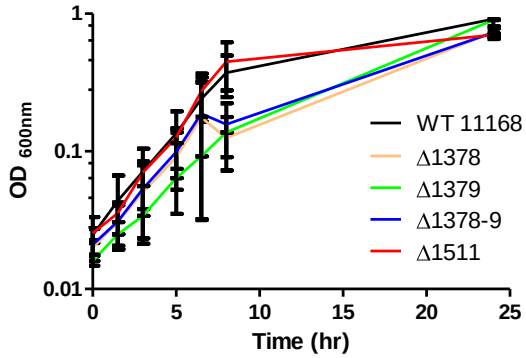
Figure 8.12 FDH activity (enzyme activity)
The FDH activity of several mutants, and chemically and genetically complemented mutants compared with the wild-type *C. jejuni* 11168. Table 8.1 identifies the abbreviations used in this figure. 'plus Sel' = 5 μM sterile selenium dioxide was added to the growth broth. The standard deviation is shown and these data are based on several biological repeats.

For completeness, mutants were made in the FDH large, selenocysteine-containing subunit gene, *fdhA*, Cj1511. This mutant was used as a negative control in the FDH assay (Figure 8.12). Δ1511 did not have detectable FDH activity, as predicted. Addition of exogenous selenium (5 μM) into the growth medium did not restore FDH activity. This result suggested that that *C. jejuni* NCTC 11168 only has one enzyme capable of FDH activity, and when this is inactivated the cells lose the ability to metabolise formate. The addition of exogenous selenium could not restore FDH activity in Δ1511, because the FDH was not present. The growth, motility and auto-agglutination phenotypes of Δ1511 were tested to ensure it was viable and any effects seen on FDH activity were due to the mutation and not a secondary effect (Figure 8.13).

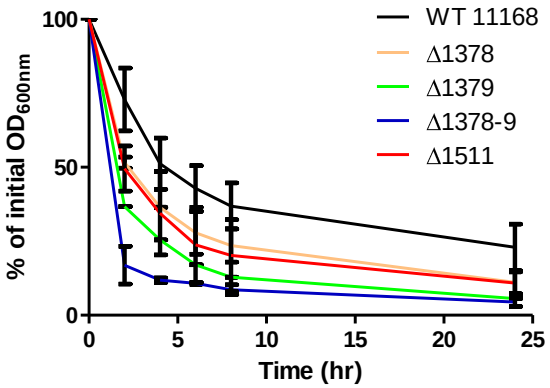
In addition to Δ1511; mutants lacking the selenium transferase gene, *selA*, Cj1378, and the putative selenocysteine-specific elongation factor gene, *selB*, Cj1379 (Δ1378, Δ1379 & Δ1378-79) were made. After establishing that these mutants had no obvious phenotype, such as growth, motility or auto-agglutination (Figure 8.13), the FDH activity of the selenocysteine synthetic defective mutants was compared with the wild-type (*C. jejuni* NCTC 11168). Δ1378, Δ1379 & Δ1378-79 all had no detectable FDH activity. The FDH activity was not restored by the addition of exogenous 5 μM selenium (Figure 8.12). This result suggests that *C. jejuni* NCTC 11168 has a single pathway for selenium incorporation, and this includes *selA* and *selB*. When the *selA* and *selB* genes were inactivated, the selenium was unable to be incorporated into the FDH. The addition of exogenous selenium could not restore FDH activity in this case, because there is no alternative pathway in *C. jejuni* NCTC 11168 for selenium incorporation into selenoproteins.

Full comparison of the mutants, complemented mutants (genetically and chemically) and *C. jejuni* wild-type strain (11168) can be seen in Figure 8.12.

a



b



c

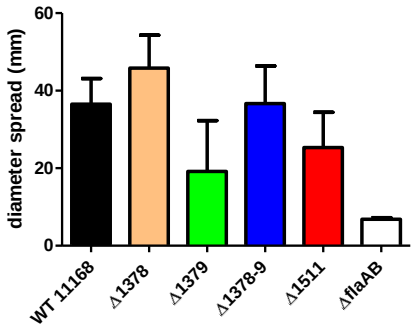


Figure 8.13 Phenotypic study of Δ1511, Δ1378, Δ1379 & Δ1378-79
Growth (a), auto-agglutination (b) and motility (c). Of the phenotypes tested, there was no significant difference between the mutants when compared to the wild-type (WT) strain *C. jejuni* NCTC 11168. This meant that the results seen in the FDH activity assay were most likely genuinely due to the mutation, and not a secondary effect derived from the mutagenesis method.

8.3.5 Nuclear Magnetic Resonance (NMR) spectroscopy revealed that slow rate formate metabolism was taking place in the Cj1501 insertion inactivation mutant, which was previously believed to be FDH-null.

C. jejuni variants (wild-type NCTC 11168, Δ1501, Δ1501::1501, Δ1511) were grown in liquid Brucella medium with and without exogenous selenium. The medium was tested for the presence of various compounds after removal of the *C. jejuni* cells. This information was used to determine the metabolism of the *C. jejuni* variants extremely precisely. In particular, the level of formate which can be accurately recorded by this NMR method is 0.05 mM.

Using an increased formate concentration in the medium (1.2 mM, compared with the normal 0.2 mM of Brucella broth), a selection of mutants, complemented mutants (genetically and chemically), and the wild-type strain were analysed over several time points for growth media component concentration (Figure 8.14). The strains used in the investigation were wild-type *C. jejuni* NCTC 11168, the insertional inactivation mutant, Δ1501, and its genetic complement, Δ1501::1501, the chemically complemented Δ1501 grown with 5 μM sterile selenium dioxide, the FDH mutant, Δ1511, with and without exogenous selenium, and a Brucella liquid medium control. The compounds monitored by NMR over several time points (0, 2, 4, 8 and 24 hours) were formate, lactate, acetate, pyruvate and succinate.

Figure 8.15 shows that, after 8 hours of growth, in Δ1501 there was formate metabolism, although it was much slower when compared to the wild-type *C. jejuni* NCTC 11168. The chemical complementation of the Cj1501 insertion inactivation mutant with exogenous selenium (Δ1501 plus Sel) increased the ability of the cells to metabolise formate. The Δ1501 plus selenium resulted in the complete depletion of formate levels after approximately 8 hours of growth. In comparison, the wild-type *C. jejuni* NCTC 11168 and genetically complemented Cj1501 inactivation mutant (Δ1501::1501) both depleted formate from the medium completely after only 2 hours of growth.

For comparison, the inability of a FDH inactivated mutant (Δ1511) to metabolise formate (with and without additional exogenous selenium) was illustrated. Formate levels remained constant in Δ1511 throughout the experiment (Figure 8.16). This suggested that *C. jejuni* NCTC 11168 does not have an additional protein capable of metabolising formate, other than FDH. Therefore, any hypotheses about formate being depleted by oxidation are unlikely and it was concluded that the depletion of formate in *C. jejuni* NCTC 11168 is due to the FDH alone.

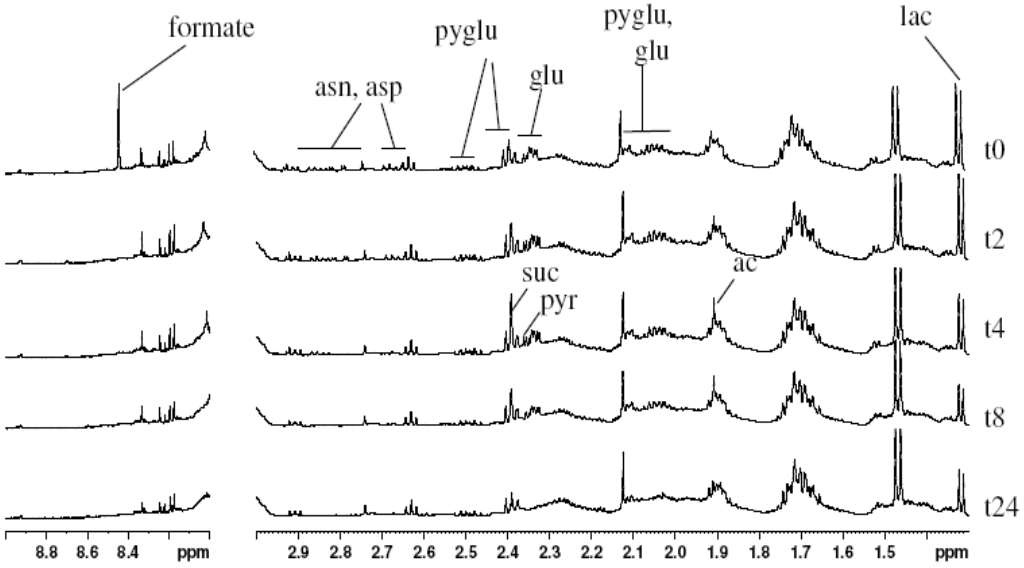


Figure 8.14 ^{1}H NMR spectra
Analysis by Gwen Le Gall.
Overlay image of ^{1}H NMR spectra of growth medium extract from wild-type *C. jejuni* NCTC 11168 taken at points (t) 0, 2, 4, 8 and 24 hours. The formate signal can clearly be seen at ~8.5 ppm and is instantly metabolised during *C. jejuni* growth. ac, acetate; lac, lactate; pyglu, pyroglutamic acid; glu, glutamic acid; suc, succinate; pyr, pyruvate; asp, aspartate; asn, asparagine.

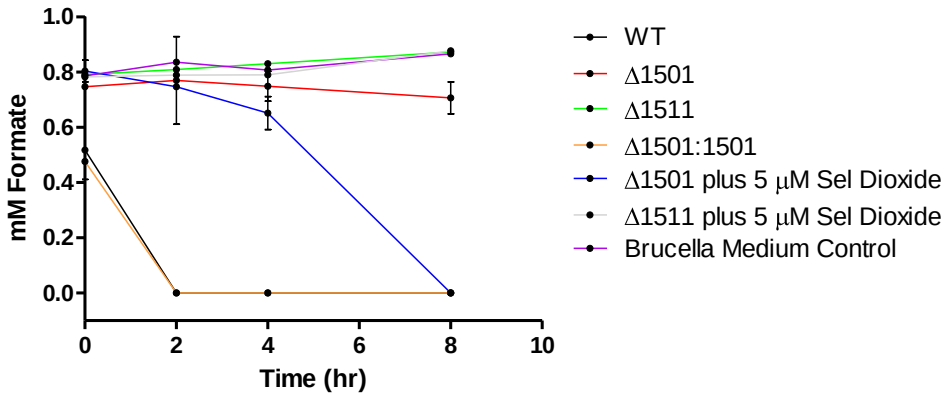


Figure 8.15 Formate metabolism in *C. jejuni*
The speed which wild-type (WT) *C. jejuni* NCTC 11168 can metabolise formate compared with Δ1501, Δ1501::1501, Δ1501 plus selenium, Δ1511 & Δ1511 plus selenium. A Brucella medium control is also shown (see legend). Δ1511 had no formate metabolism, even with the addition of exogenous selenium. The wild-type and Δ1501::1501 had rapid formate metabolism, and the Δ1501 had a slow formate metabolism, which was increased by the addition of exogenous selenium. The NMR analysis indicated that the amount of formate in the medium was 0.8 mM, whereas it was made using a known stock and calculated to be 1.2 mM. The discrepancy between the values is not noteworthy as the medium was made as a single batch and all biological repeats have the same amount of formate in the medium.

Because Δ1501 initially lacked any detectable FDH activity (as shown by anaerobic benzyl-viologen linked assay, for 200 seconds), the ability of formate to be metabolised by Δ1501 over a longer time period (8 hours) was of interest. NMR analysis clearly shows the formate levels in *C. jejuni* samples after several hours of growth (Figure 8.15). Immediately following inoculation, (0 hours) the amount of formate in both the wild-type *C. jejuni* NCTC 11168 and the complemented Cj1501 mutant (Δ1501::1501) was lower than in the other samples tested. And in both of these samples (wild-type and Δ1501::1501) the formate was metabolised completely after only two hours of growth. More interesting, the Δ1501 (which was thought, based on benzyl viologen linked assays, to be FDH-null) showed a small amount of formate metabolism after 8 hours of growth. Although the change is small, it is distinctly different from those samples which do not metabolise formate: Δ1511 and the Brucella control. The benzyl viologen linked assay showed Δ1501 plus selenium to have approximately 25% the FDH activity of the wild-type (Figure 8.12), whereas the numerical NMR data suggested that the FDH

activity of Δ1501 plus selenium to be similar to the FDH inactive mutant, Δ1511, after 2 hours of growth (Figure 8.15). It is noteworthy to understand that the NMR data is looking at formate depletion, whereas the benzyl viologen linked assay suggests the rate of enzyme activity. Although the NMR data suggests that there was no initial FDH activity in Δ1501 plus selenium, it is important to understand that the rate of FDH enzyme activity in Δ1501 plus selenium was also slower (as shown by the benzyl viologen linked assay), so the two data sets are logical. The benzyl viologen linked assay data was generated over a very short time period (200 seconds) and the method is indirect, whereas in the direct method of NMR, data was gathered at 2 hour intervals. Therefore, rate of FDH enzyme activity and the rate of formate depletion may not be exactly the same when comparing these two distinct techniques, but the trend in the data is comparable.

Figure 8.16 shows that after 24 hours of growth, the formate has been severely depleted in the Δ1501 growth medium. This suggests that although not detectable on a 2D proteomic gel, or detectable during the benzyl viologen linked assays, the inactivation of Cj1501 may not result in the complete inactivation of FDH in *C. jejuni*.

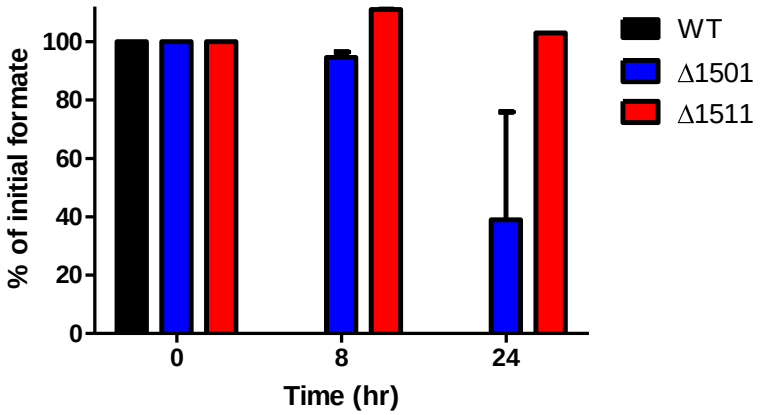


Figure 8.16 Formate metabolism in *C. jejuni*
The relative amount (%) of formate in the wild-type (WT) *C. jejuni* NCTC 11168, Δ1501 and Δ1511 growth media after 0, 8 and 24 hours of growth. The standard deviation is shown. After 24 hours, there was a range of the concentration of formate remaining between the three biological repeats of Δ1501 from complete depletion (0%) to 0.6 mM formate (80%); this explains the large standard deviation bar. Preliminary studies had suggested that formate was completely metabolised after 24 hours Δ1501 growth (data not shown).

Other metabolites (lactate, aspartate, acetate, pyruvate and succinate) were investigated to determine whether the changes seen in formate metabolism were specific to the genetic manipulations, or whether the insertional inactivation mutations were effecting the global metabolic profile of the campylobacter cells. It was determined that the effect of inactivating the gene Cj1501 was only significantly effecting formate metabolism over the 8 hour period. Figure 8.17 shows the lactate and aspartate metabolism profiles for the same samples. It can be seen that lactate and aspartate metabolism were not altered by the inactivation of either Cj1501 or Cj1511, or the complementation of the Cj1501 gene into the Cj1501-lacking mutant. Appendix III shows representative NMR trace data for the samples discussed in this section.

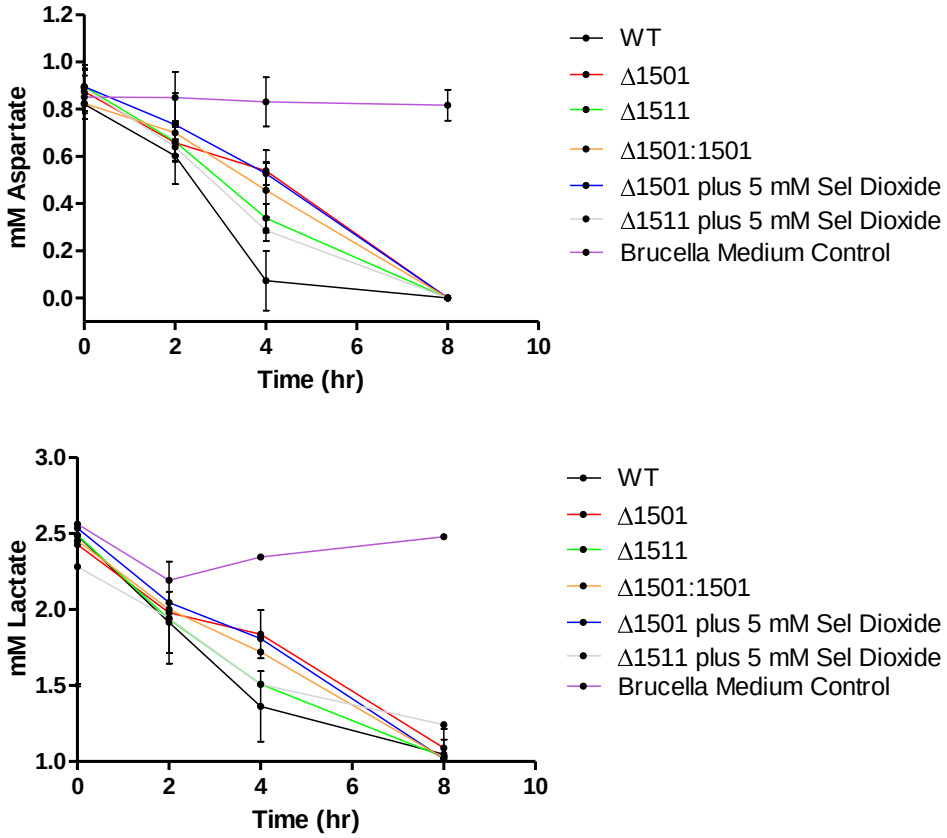


Figure 8.17 Lactate and Aspartate metabolism in *C. jejuni*
Shows the levels of aspartate (above) and lactate (below), the Δ1501, Δ1501::1501, Δ1501 plus Sel, Δ1511, Δ1511 plus Sel and wild-type(WT) *C. jejuni* NCTC 11168 strains. The NMR analysis showed the metabolism of other compounds to remain similar after genetic manipulation of the samples. The standard deviation of the NMR analysis is shown based on three biological repeats. The lactate and acetate are shown as representative metabolites.

8.3.6 Microarray analysis of *C. jejuni* with an inactivated Cj1501 showed transcriptional down-regulation of the FDH subunits

An Agilent microarray, designed at IFR to the *C. jejuni* NCTC 11168 genome, was used to examine the relative up- and down-regulation of all the genes between wild-type *C. jejuni* NCTC 11168 and Cj1501 mutant and genetic complement (described below). Thousands of probes, each specific to a gene sequence were used to determine the transcriptional changes in samples relative to the wild-type *C. jejuni*. The results are numerical and derived from the level of fluorescence bound by the labelled mRNA when hybridised to a specific gene probe. The fluorescence can be analysed numerically and interpreted statistically.

During this investigation the following samples were compared to the wild-type *C. jejuni* NCTC 11168; the insertion inactivation Cj1501 mutant (Δ1501), the genetic complement of Δ1501 (Δ1501::1501), and the wild-type strain 11168 over-expressing the Cj1501 gene, (WT::1501) (refer to Table 8.1). The comparison of wild-type with Δ1501 and wild-type with Δ1501::1501 was performed in triplicate and the comparison of wild-type with WT::1501 was performed in duplicate.

Firstly, it should be noted that the RNA used in this microarray experiment was of high quality, (run on an Agilent lab-on-a-chip and quantified by nano-drop, data not shown) and the microarray data was of sufficient quality for robust statistical analysis. Preliminary analysis indicated that the array had been successful because the correct antibiotic resistance cassettes were highlighted as up-regulated in the correct strains.

Statistical analysis programs: Firstly, the data was analysed using an excel macro [as described in 15], this macro will be defined as 'Marray' throughout this thesis. The application was used to lowess normalise (to remove gradient effects) the data and the technique 'significance analysis of microarrays' was used to determine which genes were significantly up- or down-regulated. The Marray program is an appropriate system to analyse two-state biological comparisons, such as these: wild-type Vs Δ1501, wild-type Vs Δ1501::1501,

wild-type Vs WT::1501. Secondly, the Rank Products method [16] was applied to the wild-type Vs Δ1501 and wild-type Vs Δ1501::1501 data sets. The over-expressing wild-type data (wild-type Vs WT::1501) could not be analysed by this method as three repeat data sets were required for the analysis. The Rank Products method is based on calculating rank products from replicate experiments. It provides a straightforward and statistically stringent way to determine the significance level for each gene and allows for the flexible control of the false-detection rate. The Rank Product method was designed to reduce any false-positives whilst still being able to examine a large number of genes simultaneously.

Major results from microarray statistical analysis: The repeatable and significant result, after Rank Product and Marray analysis, of this experiment was that in the Δ1501, the genes responsible for FDH subunits (Cj1509c, Cj1510c & Cj1511c) were all down-regulated (Table 8.2). Each of the three genes contained several probes, and each probe was present twice on the microarray slide. The scatter plots show the FDH gene probes (Cj1509c, Cj1510c & Cj1511c) highlighted in yellow for all three biological repeat data sets and the data trend is clearly visible (Figure 8.18). The plot is generated by plotting the relative green and red fluoresce of each probe on two logarithmic scales. The unaltered data appears in a linear fashion, whereas those probes which vary from the wild-type cluster on the extremity of the data set when represented in this way.

The down-regulation of Cj1511c, Cj1510c and Cj1509c suggests that when the Cj1501 is inactivated, the FDH gene is not fully transcribed. Unfortunately, it cannot be determined if transcription is halted at the UGA codon (selenocysteine), because the microarray does not contain a probe within Cj1511c and upstream of the UGA (Figure 8.19 and Appendix IV). It would be interesting to determine if the transcription was halted at the UGA because it is possible that Cj1501 has a crucial role in ensuring the selenocysteine is able to be formed and entered into the amino acid sequence. However, it is likely that although the RNA required for FDH is formed, it is not translated past the

UGA codon and is therefore unstable, and either lost or degraded by the *C. jejuni*.

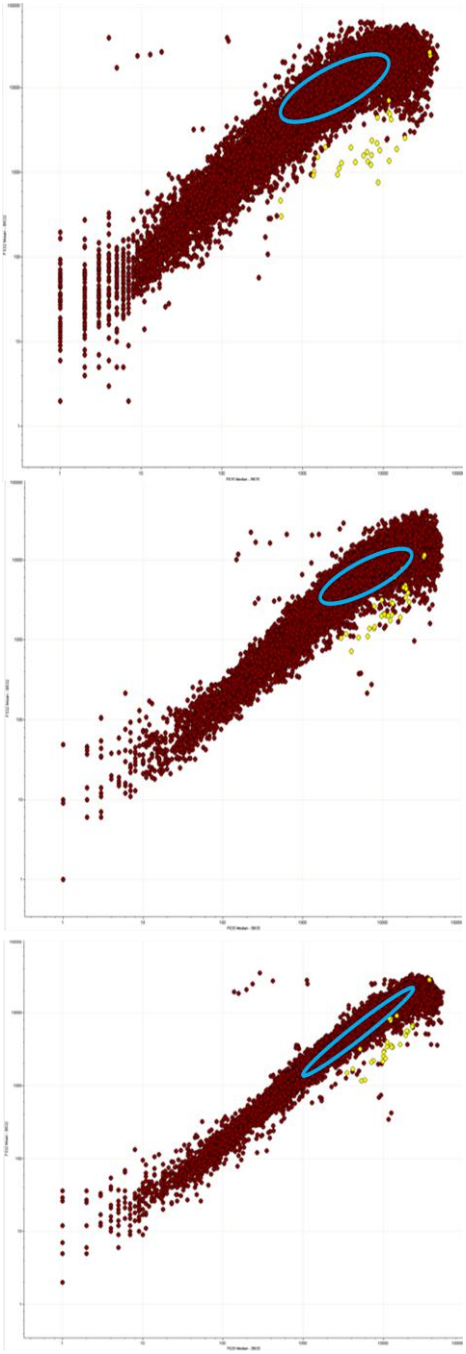


Figure 8.18 Microarray probes shown diagrammatically
Image generated using GenePixPro
The down regulation of all the probes for genes Cj1511c, *fdhA*, Cj1510c, *fdhB* and Cj1509c, *fdhC* (yellow), in the three biological repeats when the Δ1501 transcriptome was compared with the wild-type. All the other probes are shown as maroon. Probes for Cj1514c and Cj1513c, the preceding genes to Cj1511c, are shown in the area depicted with the blue ovals (0): no significant change in transcriptional regulation was observed in those probes.

Gene	Δ1501	Δ1501:1501	WT:1501
	Fold change when compared to the wild-type		
fdhM Cj1514c	1.5 ± 0.1	2.6 ± 1.7*	1.8 ± 0.6*
Cj1513c	1.8 ± 0.8	2.4 ± 1.7*	1.9 ± 0.8*
fdhA Cj1511c	**-4.2 ± 2.2**	1.2 ± 0.2*	1.1 ± 0.0*
fdhB Cj1510c	**-8.6 ± 7.1**	2.4 ± 0.8*	1.6 ± 0.0*
fdhC Cj1509c	**-6.6 ±5.2**	3.7 ± 1.5*	2.7 ± 1.3*

Table 8.2 Fold regulation of the genes in the FDH operon compared with the wild-type.
Standard deviation shown, based on three (* two) biological repeats. The data was normalised using the Marray program. The down regulated genes (*fdhA-C*) are shown in **bold**. *fdhA-C* were down regulated in all three biological repeats, when compared to the wild-type, between two and 18-fold.

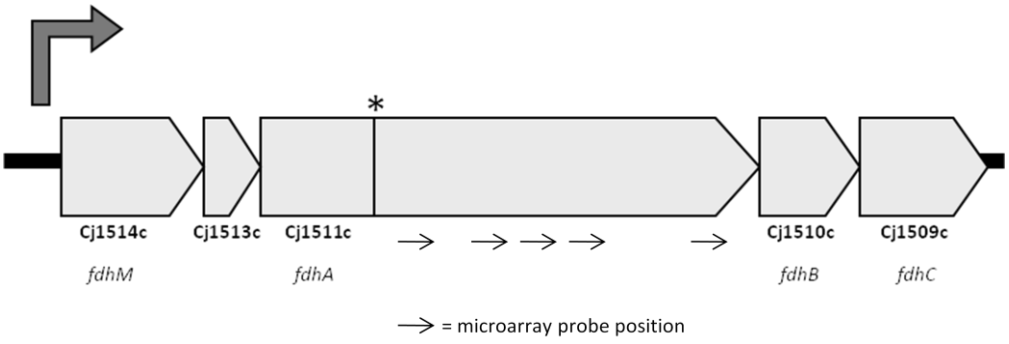


Figure 8.19 A diagrammatic representation of the FDH operon
* represents the location of the selenocysteine. The probe positions are depicted with arrows, the probe sequences are given in Appendix IV.

8.3.7 Purification of Cj1501

The small gene Cj1501 was amplified using specific primers from *C. jejuni* NCTC 11168 genomic DNA, and expressed in an *Escherichia coli* expression host (DE3)BL21 and pET21a vector system. After auto-inducing the *E. coli*, the over-expressed protein was purified using a T7 affinity column and run on a 4 - 12 % SDS Nu-page gel (Figure 8.20). The purified recombinant protein was buffer exchanged for long-term storage at -80°C. The protein is ready for further experiments.

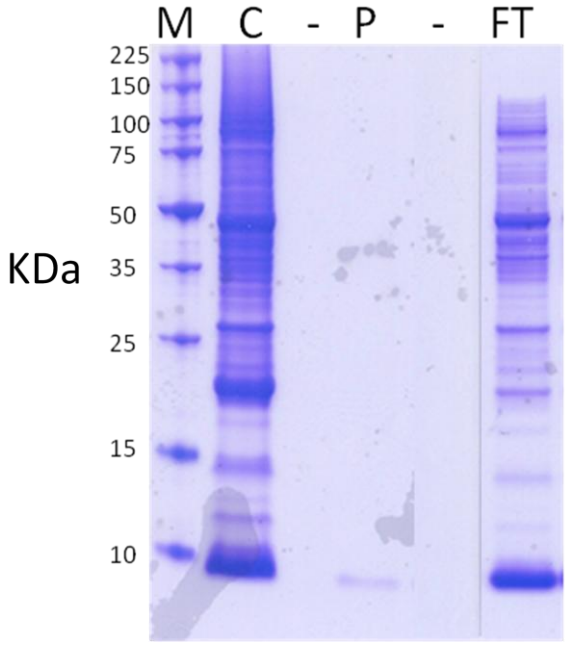


Figure 8.20 SDS page gel of Cj1501
The protein gel showing purified Cj1501 protein of approximate size 8KDa (P). M = marker (Promega), C = crude cellular protein extract of induced *E. coli*, - = empty lane, FT = flow through from T7 column.

8.4 Discussion

Does *C. jejuni* gene Cj1501 have a role in *C. jejuni* acid resistance?

It is known that most bacteria have mechanisms to allow survival at pHs either above or below their optimum [17, 18]. Elucidation of the response *C. jejuni* undertakes when subjected to a low pH could have interesting applications in food safety industry and in the understanding of *C. jejuni* pathogenesis.

When *C. jejuni* NCTC 11168 lacked the genes Cj1500 and/or Cj1501, there was no differential growth, motility or auto-agglutination phenotype when compared with the wild-type strain. During this investigation, it was not established if the insertion inactivation of Cj1501 (Δ1501) was more robust at surviving acid-shock compared to the wild-type. Previous transcriptional investigation showed that the Cj1501 gene was up-regulated on acid-shock. This result suggested that Cj1501 had a role in surviving acid-shock, and perhaps a mutant lacking Cj1501 would be more sensitive to acid conditions. However, during this investigation it was not established if Cj1501 is essential for *C. jejuni* survival at acidic pH. This would be an interesting area to research further, perhaps using more sensitive techniques. It was hypothesised that a strain lacking Cj1501 (Δ1501) would be more acid sensitive than the wild-type *C. jejuni*, because if Cj1501 was constantly up-regulated upon acid-shock, it is possible it has a role in acid-resistance, or *C. jejuni* survival. For future experimentation, sensitive pH equipment would be necessary as this investigation concluded there was a variation in *C. jejuni* survival between pH 3 – 4. At the cusp of *C. jejuni* survival (pH 3.5), the results between biological repeats were inconsistent. One explanation for this is that the media could change between 3.4 – 3.6 after inoculation, or during short term storage, and these slight, perhaps even undetectable, changes in pH may have been significant enough to cause the cells to either die or survive. An improved method would test the survival ability of *C. jejuni* 0.1 increments. Previous studies involved testing at pHs between 0.5 increments, and this was not sensitive enough to determine if there was a phenotypic difference between mutant Cj1501 and wild-type strains. This investigation showed

that *C. jejuni* cells are able to survive until approximately pH 3.5, a more acidic pH resulted in complete loss of viability. Constant pH monitoring of the growth media would be a useful tool in understanding whether the *C. jejuni* are able to lower the pH of the environment during cellular respiration.

Because *C. jejuni* NCTC 11168 without the Cj1500 and/or Cj1501 genes had no differential growth, motility or auto-agglutination phenotype observed when compared to the wild-type. It was concluded that these genes are not necessary for these functions (growth, motility and auto-agglutination). And consequentially, the genes (Cj1500 and Cj1501) are most likely not required for the regulation of biochemical pathways controlling aspects of growth, motility or auto-agglutination, such as flagella assembly, structural protein synthesis and cell division. This was not a strain specific phenotype, as the same results were observed when the Cj1500-01 homologue was inactivated in *C. jejuni* strain 81116. To be able to determine that a gene does not a have a certain role, is just as important as determining what the role is.

The relationship between the *C. jejuni* genes Cj1500 and Cj1501 and the FDH enzyme (Cj1509-11c). Using several techniques, including proteomics, transcriptomics, metabolomics and enzyme assays, it has been concluded that the *C. jejuni* NCTC 11168 gene Cj1501 is essential for the production of a fully functioning FDH enzyme. This result was interesting because a role for the Cj1501 has never been established, or predicted, previously. This was the first example of a Cj1501 homologue having a role regarding active FDH production.

The insertion inactivation mutants of these two sequential genes (Δ1501, Δ1500, Δ1500-01) all lacked detectable formate dehydrogenase (FDH) activity when live cells were tested in a benzyl viologen linked assay. The FDH enzyme consists of several subunits (Cj1511c, *fdhA,* Cj1510c, *fdhB,* and Cj1509c, *fdhC*). FDH has been reported as being important in *C. jejuni* colonisation of chickens [12], and it is the only known selenoprotein in *C. jejuni*. Interestingly, Cj1500 and Cj1501, or their homologues in different species, have never been linked to FDH or selenoprotein activity previously.

The genes of interest here (Cj1500-01 and Cj1509c-11c) are located in a similar region of the *C. jejuni* NCTC 11168 chromosome, however they are not related to one another because they are located on opposite strands. It is highly unlikely that the mutagenesis of Cj1500-01 resulted in the inactivation of Cj1509c-11c in anyway.

Proteomics and FDH Assay Discussion: When the proteome of *C. jejuni* lacking Cj1501 (Δ1501) was compared to that of the wild-type; the FDH subunits *fdhA* and *fdhB* were missing in the mutant (identified by 2D separation and confirmed by mass spectroscopy), but present in the wild-type. To date, *fdhC* has not been identified in *C. jejuni* NCTC 11168 using proteomics. Further to this, living mutant cells (Δ1501, Δ1500 & Δ1500-01) did not have detectable FDH activity when assayed anaerobically using the benzyl viologen linked assay. It was concluded that Cj1501 protein has a crucial, but still undefined, role in FDH formation and activity. The FDH enzyme was not present in the Δ1501 (confirmed by proteomics) and the cells have no detectable FDH activity (confirmed by FDH enzyme assay). However, the FDH activity could be restored in two ways; firstly, the re-expression of the Cj1501 gene under a natural *C. jejuni* promoter (Δ1501::1501) resulted in partial restoration of FDH activity, and secondly, growing the Cj1501 mutant cells in media containing exogenous selenium partially restored the FDH activity (Figure 8.12). When a mutant lacking a specific gene has a certain phenotype, and is forced to re-express that gene and the phenotype returns to that of the wild-type, it can be said that the gene in question is responsible for the phenotype. Here, it is understood that the Cj1501 gene is definitely responsible for the functioning FDH in *C. jejuni*. However, in addition to genetic complementation, FDH activity could also be restored by the addition of exogenous selenium. This information allows the further definition of the role the Cj1501 protein is performing in *C. jejuni*. Whatever process is not being completed in the *C. jejuni* cells due to the inactivation of the Cj1501 gene, the process can be restored by the addition of exogenous selenium. It is possible that perhaps an alternative pathway, or transporter, necessary for a fully functional FDH was being activated due to the high selenium concentrations.

It cannot be concluded, based solely on the experiments in this investigation, that the exogenous selenium restores the FDH protein complex in Δ1501. The assay shows the restoration of FDH activity, however, until the cellular proteins of Δ1501 grown in a selenium supplemented medium are examined, the selenium could be responsible for the induction of another FDH, or another dehydrogenase enzyme, which has the ability to metabolise formate. 2D proteomics would establish whether the known FDH (Cj1511-09c) is being reconstituted in the Δ1501 cells grown in selenium supplemented medium.

In addition to Δ1501 lacking FDH activity, Δ1500 & Δ1500-01 also lacked FDH activity when living cells were assayed. Proteomics was not performed on Δ1500 & Δ1500-01, but it is likely that they too did not contain the FDH enzyme. The exogenous addition of selenium to the living *C. jejuni* cells Δ1500 & Δ1500-01 resulted in partial restoration of FDH activity (Figure 8.12). (This was the same effect as seen in Δ1501.) The Cj1500 protein is annotated as containing transmembrane domains on the Pfam website (http://pfam.sanger.ac.uk/). It is possible that Cj1500 is a transmembrane transporter of selenium, and inactivation of this transmembrane protein results in an inactive FDH, because the selenium is not able to enter the cell, and therefore enter the selenoprotein. Cj1501 may have a role as an intracellular transporter of selenium. It was hypothesised that the paired function of Cj1500 and Cj1501 was a selenium uptake system. In the absence of Cj1500 and Cj1501 and an excess of selenium, the activity of FDH was restored. This suggests that selenium is being incorporated into the FDH, via the selenocysteine amino acid, to create a fully functioning FDH. It is possible that in cases of extreme concentrations of selenium, an alternative uptake system was used in Δ1501, Δ1500 & Δ1500-01 and this resulted in incorporation of a selenocysteine into the FDH protein.

To determine more about the role of the Cj1500 gene, the Cj1501 gene was expressed in Δ1500 & Δ1500-01 (named Δ1500::1501 & Δ1500-01::1501). This experiment was necessary to determine that in Δ1500, the only gene inactivated was Cj1500, and that the insertion cassette had not had an effect

downstream which resulted in both Cj1500 and Cj1501 being inactivated. However, unlike Δ1501::1501, FDH activity was not restored in both Δ1500::1501 & Δ1500-01::1501. This suggested that Cj1500 was also essential in the production in a functioning FDH.

This FDH-lacking phenotype is not a strain specific phenomenon as it was also shown in *C. jejuni* 81116. Interestingly, the pair of sequential genes Cj1500 and Cj1501 from *C. jejuni* NCTC 11168 appear as a single gene in *C. jejuni* 81116; C8J_1404. Homologues of Cj1500 and Cj1501 are often found in close proximity to one another across several bacterial species. However, C8J_1404 is the only known example of the genes fused together as a single open reading frame. An insertion mutant was made in the Cj1500-01 homologue; C8_J1404 (Δ1404). The mutant showed the same FDH-null phenotype as the Δ1500, Δ1501 & Δ1500-01, when assayed anaerobically using the benzyl viologen-linked assay.

Future possible genetic complementation experiments: To further investigate the physiology of the two sequential genes, Cj1500-01, the single mutant Δ1500 and double mutant Δ1500-01 could be complemented with the inactivated genes, Cj1500 and Cj1500-01 respectively. A complementation such as this theoretically would result in FDH activity being restored. This result would undoubtedly conclude that the genes Cj1500-01 are responsible for a fully active FDH enzyme.

To investigate the role of the Cj1500-01 homologues across various species, the mutants (Δ1501, Δ1500 & Δ1500-01) could be complemented with homologues from another species, such as the β-proteobacterium *N. europaea*, a Gram-negative obligate chemolithoautotroph, and the hyperthermophillic bacterium *T. petrophila*, which both contain Cj1500-01 homologues positioned adjacently in the genome, and determine if the FDH activity can be restored in *C. jejuni*. These genes are highly conserved across several species, so it is possible that the application of a non-native gene into the Cj1500-01 inactivated mutant would restore FDH activity. However, due to the low GC nature of the *C. jejuni* genome, careful consideration of which homologue to

use would need to be considered. Codon usage is essential to consider when expressing a non-native gene in an organism; the gene should require amino acids which are available in the host organism. This experiment would provide a physical role, to those genes currently sequenced, but un-annotated, and could provide insight into the extent of this genetic conservation across species. These genes Cj1500-01, and their homologues, have never been reported upon in so much detail before.

Is *C. jejuni* FDH a selenoprotein? This investigation has confirmed that FDH is most likely a selenoprotein, because mutants lacking the selenium transferase, *selA*, Cj1378, and the putative selenocysteine-specific elongation factor, *selB*, Cj1379, individually and combined (Δ1378, Δ1379 & Δ1378-79) had no detectable FDH activity. When exogenous selenium is added to the media, the FDH activity is not restored. This suggests that the Cj1378 and Cj1379 genes ensure the insertion of a selenium (in the form of selenocysteine) to the FDH. An excess of exogenous selenium could not restore FDH activity in *C. jejuni* with the *selA* and *selB* genes inactivated. Therefore, it can be concluded that there are no alternative back-up pathways, genes or mechanisms for selenium incorporation into a selenocysteine in *C. jejuni* NCTC 11168 other than *selA* and *selB*.

Metabolomic Data Discussion: NMR of the growth medium which contained the Cj1501 insertion inactivation mutant (Δ1501) was tested metabolically after several hours of growth and the metabolites present were compared with wild-type *C. jejuni* NCTC 11168 strain and the *C. jejuni* strain lacking the FDH large subunit (*fdhA*, Δ1511). NMR is a highly sensitive technique used for metabolite detection, such as formate. Within the NMR trace, the formate peak is isolated without any interfering groups; therefore the level of detection can be as low as approximately 0.05 mM formate. Because Brucella broth is a rich medium, it is not always possible to detect concentrations as low as this for all *C. jejuni* metabolites using NMR.

It was concluded, using NMR, that *C. jejuni* NCTC 11168 only has one enzyme capable of metabolising formate. It was seen that the inactivation of *fdhA,* the

FDH large subunit, led to the inability of the cell to metabolise formate, as formate levels in the growth medium remained consistent for several hours after inoculation.

Although the benzyl viologen linked assay suggested that the inactivation of Cj1501 (Δ1501) led to an inactivation in FDH, it can be seen using NMR, a more sensitive technique, that formate is actually being metabolised, but at a rate much lower than that of the wild-type. After 24 hours of Δ1501 growth, the amount of formate had decreased to approximately 30% (on average) of the original concentration. This suggested that the Δ1501 mutant does have a limited ability to metabolise formate. Perhaps the Δ1501 mutant does have an active FDH enzyme, but the concentration is so low that it was not detected on the 2D proteomic gel, and the activity below the level of detection in the benzyl viologen linked assay. If Cj1500 and Cj1501 are functioning as a selenium uptake system, and this system is inactivated, perhaps an alternative uptake system, or a specific channel, which has a very low affinity for selenium is functioning. This alternative system may have moved selenium intracellularly and this resulted in an active FDH enzyme. It is known that the reduction in formate concentration must be due to the FDH, because when the FDH was inactivated (by insertion inactivation of *fdhA*, Δ1511) there was no formate metabolism. Assuming that the Δ1501 mutant has a completely inactivated Cj1501 by the mutagenesis process, there must be another system in *C. jejuni* capable of fulfilling the role of Cj1501 and ultimately aiding FDH function.

Genome Regulation Discussion and Interpretation: It has been established, that when the Cj1501 of *C. jejuni* NCTC 11168 is inactivated by an insertion cassette, FDH activity is not detected in the living cells, and in addition, proteomic analysis shows the FDH protein to be absent. However, NMR analysis suggests a small background level of FDH activity. Analysis at the transcriptional level (by whole genome microarray) showed that when Cj1501 was inactivated, the level of RNA for the FDH major subunit genes (*fdhA, fdhB* and *fdhC*) was significantly lower than in the wild-type. The relative amount of RNA detected in the Δ1501 mutant was between two and

18-fold less than the wild-type *C. jejuni* RNA concentration for all three FDH protein subunits in all three biological repeats.

Transcript start site analysis performed at the IFR (data not published) suggested that the mRNA for the FDH gene starts at Cj1514c. It can be ruled out that transcriptional regulation is responsible for the lack of FDH RNA, because genomic analysis indicated that the promoter of the FDH operon precedes Cj1514c and lies two genes upstream of the first FDH subunit (*fdhA*, Cj1511c, refer to Figure 8.19). If the operon was being transcriptionally regulated, then the whole operon would be down-regulated. A down-regulation in the genes Cj1514c and Cj1513c was not seen, in contrast to that of Cj1511c, Cj1510c and Cj1509c (*fdhA, fdhB* and *fdhC*), in the Δ1501 mutant.

It is more likely that the down-regulation of *fdhA, fdhB* and *fdhC* is due to post-transcriptional modification. If post-transcriptional regulation takes place, the control of gene expression occurs between the transcription and translation of the gene. It is possible that when the mRNA was being translated, the presence of the selenocysteine codon (UGA) caused translation to stall, resulting in destablisation of the mRNA (Figure 8.21). It is known that when bacterial cells are grown without adequate selenium, translation of the selenoprotein may be terminated at the UGA [14]. This provided a possible explanation as to why the preceding genes in the operon (Cj1514-13c) were not down-regulated and why the FDH protein was not being produced (as confirmed by 2D proteomics).

2D proteomics did not reveal a truncated FDH protein, however, if it was present it would have been fairly small at 181 amino acids, the predicted mass of the protein is 20 KDa, with a pI 8.94. And, if it were at low concentrations, it may have been below the level of detection of the proteomics equipment and methods used. Normally, the proteomics method can visualise most *C. jejuni* intracellular and periplasmic proteins. However, it is limited by concentration; a large volume of a single protein, e.g. major outer membrane protein, can dominate the gel image.

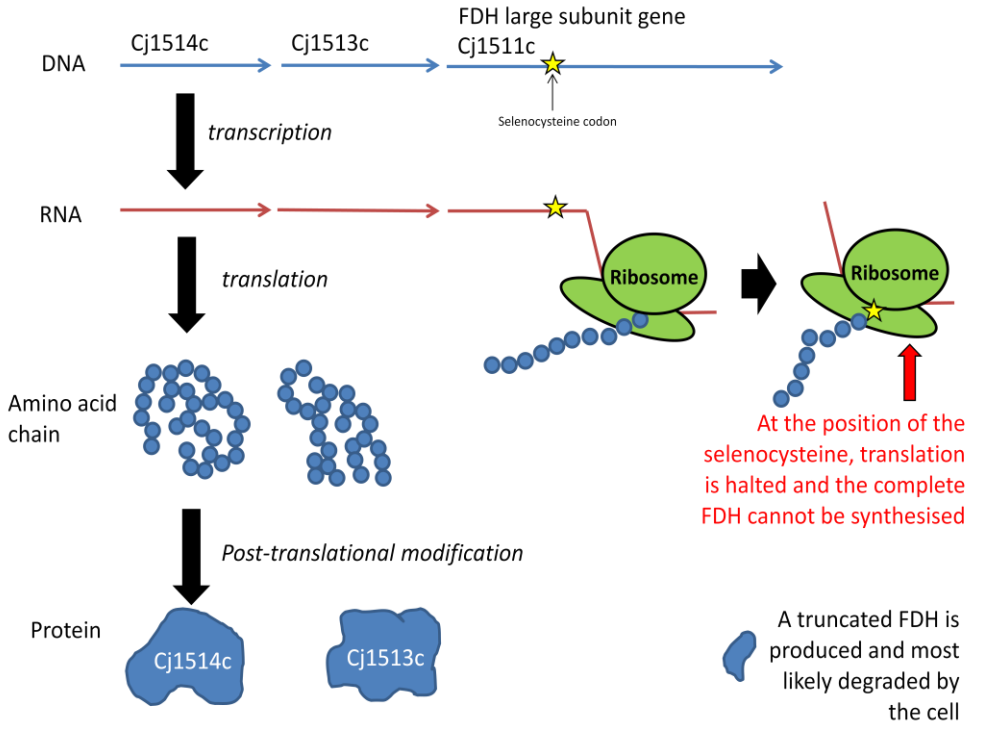


Figure 8.21 Shows the possible transcription and translation of the FDH operon

The inactivation, or over-expression, of Cj1501 in *C. jejuni* NCTC 11168 did not result in the consistent up- or down-regulation of any additional genes (except the FDH synthetic genes, Cj1511-09c, as previously discussed) when examined by microarray. This suggests that Cj1501 is only involved in the FDH production pathway, because the rest of the genome remained insignificantly affected. It is hypothesised that Cj1501 is involved with the formation (via the selenocysteine production route) of other, currently undiscovered, selenoproteins in *C. jejuni*, and a microarray experiment would highlight those unknown selenoproteins. However, because no other gene groups were significantly regulated by the lack or over expression of Cj1501, it is likely that there is not another selenoprotein in *C. jejuni* which requires Cj1501 for normal formation. Currently, there are no other known selenoproteins in *C. jejuni*, therefore the lack of Cj1501 influenced genes agrees with the suggestion of no further selenoproteins in the organism.

Future transcriptomic experiments: Further possible investigations to determine more about the regulation of the FDH include transcriptomic

investigation of differentially mutated strains of *C. jejuni*, such as Δ1511, which has the Cj1511 (*fdhA*) inactivated by the insertion of an antibiotic cassette. Without a functioning FDH, it would be interesting to determine if additional genes are differentially regulated in compensation for the inactivation of FDH. Further to this, transcriptome and proteome investigation of *C. jejuni* with inactivated Cj1500 and/or Cj1501 grown in a selenium, or selenocysteine, supplemented medium would be an interesting investigation. If the lack of selenocysteine is causing the apparent down-regulation of the FDH genes (Cj1511-09c) in Δ1501, then it would be interesting to determine if genomic regulation is returned to normal in the presence of a selenocysteine. Investigations such as these could lead to a further understanding of how the synthesis of FDH in *C. jejuni* is controlled and regulated. In addition, systems such as this may in place in a variety of micro-organisms, and this knowledge could be applied further afield than the Campylobacter species.

Future experiments to elucidate the function of the Cj1501 gene: To further investigate the specific role of Cj1501, the purified recombinant protein could be tested for properties, such as the ability to bind a selenium ion. Mass spectroscopy of the protein before and after a selenium incubation may answer this question.

Another, more advanced technique to elucidate the role of Cj1501 would be to define the structure of the protein using x-ray crystallography. This may lead to the discovery of a structure specific role of the protein, such as the ability to bind another molecule or polymerise. If the crystal structure can be determined, the understanding of the protein function will be enhanced beyond the understanding based on amino acid sequence alone. For example, how the protein interacts with other molecules, possibly selenium, and how, and if, it undergoes conformational changes can be studied using crystallography. As minimal investigations have been performed on this protein, any insight would be helpful in our understanding of Campylobacter and other micro-organisms.

Biochemical investigation of the Cj1501 protein would be the next step in this study. Experiments into whether the small protein is able to bind selenium could be performed using the purified protein. Mass spectroscopy of protein with, and without, selenium incubation could be compared. If the protein is able to bind selenium, the molecular weight would increase as selenium has a relative atomic mass of 78. Obviously, there are issues associated with this technique, and because a protein like this has not been investigated in this way previously the parameters under which to test the protein would have to established. Firstly, the protein would need to be stripped of all co-factors and possible metals which may bind to the surface. Typically, this can be done by a series of pH controlled washes. Then, it would be necessary to develop the optimum selenium binding incubation conditions; which may be temperature or oxygen concentration dependent.

Concluding Remarks

This investigation has provided some interesting findings regarding the small, highly conserved *C. jejuni* gene Cj1501, and the preceding gene, Cj1500. In this thesis the role of this small gene has been examined, and a link between the genes and FDH activity has been established. The future expansion of this investigation will lead to a greater understanding of the gene function in *C. jejuni*. The unique regulation, presence and expression of certain genes can be linked to *C. jejuni* survival and pathogenicity.

Chapter Eight – References

1. Janssen R, Krogfelt KA, Cawthraw SA, van Pelt W, Wagenaar JA, Owen RJ: **Host-pathogen interactions in Campylobacter infections: the host perspective**. *Clinical Microbiology Reviews* 2008, **21**(3):505-518.
2. Young KT, Davis LM, Dirita VJ: ***Campylobacter jejuni*: molecular biology and pathogenesis**. *Nature Reviews Microbiology* 2007, **5**(9):665-679.
3. Teplitski M, Goodier RI, Ahmer BM: **Pathways leading from BarA/SirA to motility and virulence gene expression in Salmonella**. *Journal of Bacteriology* 2003, **185**(24):7257-7265.
4. Pearson BM, Gaskin DJ, Segers RP, Wells JM, Nuijten PJ, van Vliet AH: **The complete genome sequence of *Campylobacter jejuni* strain 81116 (NCTC11828)**. *Journal of Bacteriology* 2007, **189**(22):8402-8403.
5. Hazeleger WC, Wouters JA, Rombouts FM, Abee T: **Physiological activity of *Campylobacter jejuni* far below the minimal growth temperature**. *Applied and Environmental Microbiology* 1998, **64**(10):3917-3922.
6. Chiang SL, Taylor RK, Koomey M, Mekalanos JJ: **Single amino acid substitutions in the N-terminus of *Vibrio cholerae* TcpA affect colonization, autoagglutination, and serum resistance**. *Molecular Microbiology* 1995, **17**(6):1133-1142.
7. Knutton S, Shaw RK, Anantha RP, Donnenberg MS, Zorgani AA: **The type IV bundle-forming pilus of enteropathogenic *Escherichia coli* undergoes dramatic alterations in structure associated with bacterial adherence, aggregation and dispersal**. *Molecular Microbiology* 1999, **33**(3):499-509.
8. Roggenkamp A, Neuberger HR, Flugel A, Schmoll T, Heesemann J: **Substitution of two histidine residues in YadA protein of *Yersinia enterocolitica* abrogates collagen binding, cell adherence and mouse virulence**. *Molecular Microbiology* 1995, **16**(6):1207-1219.

9. Guerry P: **Campylobacter flagella: not just for motility**. *Trends in Microbiology* 2007, **15**(10):456-461.
10. Golden NJ, Acheson DW: **Identification of motility and autoagglutination *Campylobacter jejuni* mutants by random transposon mutagenesis**. *Infection Immunity* 2002, **70**(4):1761-1771.
11. Kelana LC, Griffiths MW: **Use of an autobioluminescent *Campylobacter jejuni* to monitor cell survival as a function of temperature, pH, and sodium chloride**. *Journal of Food Protection* 2003, **66**(11):2032-2037.
12. Weerakoon DR, Borden NJ, Goodson CM, Grimes J, Olson JW: **The role of respiratory donor enzymes in *Campylobacter jejuni* host colonization and physiology**. *Microbial Pathogenesis* 2009, **47**(1):8-15.
13. Smart JP, Cliff MJ, Kelly DJ: **A role for tungsten in the biology of *Campylobacter jejuni*: tungstate stimulates formate dehydrogenase activity and is transported via an ultra-high affinity ABC system distinct from the molybdate transporter**. *Molecular Microbiology* 2009, **74**(3):742-757.
14. Arner ES: **Recombinant expression of mammalian selenocysteine-containing thioredoxin reductase and other selenoproteins in *Escherichia coli***. *Methods in Enzymology* 2002, **347**:226-235.
15. Holmes K, Mulholland F, Pearson BM, Pin C, McNicholl-Kennedy J, Ketley JM, Wells JM: ***Campylobacter jejuni* gene expression in response to iron limitation and the role of Fur**. *Microbiology* 2005, **151**(Pt 1):243-257.
16. Breitling R, Armengaud P, Amtmann A, Herzyk P: **Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments**. *Febs Letters* 2004, **573**(1-3):83-92.
17. Cotter PD, Hill C: **Surviving the acid test: responses of gram-positive bacteria to low pH**. *Microbiol Mol Biol Rev* 2003, **67**(3):429-453, table of contents.

18. Audia JP, Webb CC, Foster JW: **Breaking through the acid barrier: an orchestrated response to proton stress by enteric bacteria**. *Int J Med Microbiol* 2001, **291**(2):97-106.

Chapter Nine

Bioinformatic analysis of the *C. jejuni* genes: Cj1500 and Cj1501

TABLE OF TABLES

TABLE OF FIGURES

Chapter Nine - Bioinformatic analysis of the *C. jejuni* genes: Cj1500 and Cj1501

9.1 Introduction

The small gene Cj1501 from the *Campylobacter jejuni* NCTC 11168 genome was identified as a gene of particular interest during a microarry experiment which indicated this gene was consistently up-regulated in response to different acid-shock conditions (Figure 8.2). Further investigation of the Cj1501 gene concluded that the small gene had a crucial role in the formation of formate dehydrogenase (FDH), a selenocysteine-containing tungsten-binding enzyme (see Chapter Eight for background information and full experimental details). Genes which are up-regulated in acid-shock may be important for *C. jejuni* infection and pathogenicity. FDH has been shown to have an important role in optimum *C. jejuni* colonisation of the chicken caecum [1]. Therefore, elucidation of the role of Cj1501 may have important implications in the understanding of *C. jejuni* pathogenicity, virulence, survival and infection.

The String database (http://string-db.org/ [2]) and the Basic Local Alignment Search Tool (BLAST - http://blast.ncbi.nlm.nih.gov/Blast.cgi [3]) indicated that the small gene, Cj1501, was highly conserved in Campylobacter species, and homologues were found in phylogenetically diverse organisms, such as the β-proteobacterium *Nitrosomonas europaea*, a Gram-negative obligate chemolithoautotroph, and the hyperthermophillic bacterium *Thermotoga petrophila*. Also of interest was the conserved sequential distribution of genes Cj1500 and Cj1501 across species. The String database is a collection of known and predicted protein interactions. These interactions include direct (physical) and indirect (functional) associations. The interactions are derived in four ways; genomic context, high-throughput experiments, co-expression data and previous knowledge. The database currently covers 2,590,259 proteins from 630 organisms. The BLAST tool finds regions of local similarity between sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. BLAST can be used to infer functional and evolutionary relationships between

sequences as well as help identify members of gene families. This conservation of the gene (Cj1501), and preceding gene (Cj1500), across species indicated that the genes may be acting in a coupled or linked function, and could be performing a crucial role in the different bacterial species.

9.2 Objectives & Hypotheses

A collection of hypotheses were determined based on the biochemical data regarding Cj1501. When the Cj1501 gene was inactivated in *C. jejuni* NCTC 11168 by insertion of an antibiotic cassette, there was an substantial decrease of FDH activity. Addition of exogenous selenium to the Cj1501-lacking *C. jejuni* resulted in FDH activity. Therefore, perhaps Cj1501 has a role regarding selenium incorporation into FDH, which is a selenoprotein. The same result was seen when investigating a *C. jejuni* with an inactive Cj1500. Presently, there is minimal bioinformatic information on both Cj1500 and Cj1501, however, preliminary investigations showed that the two genes were conserved across the Campylobacter species and within some more phylogenetically diverse micro-organisms. This suggests that the genes have crucial role in *C. jejuni* and elucidation of the role could provide new information into the biology of the Campylobacter species.

The hypotheses and questions investigated bioinformatically in this chapter are summarised below:

- Are Cj1501 and Cj1500 consistently found together across species phylogenetically diverse from *C. jejuni*?

- Cj1500 and Cj1501 are only found in organisms containing selenoproteins

- Cj1501 is a selenium binding protein

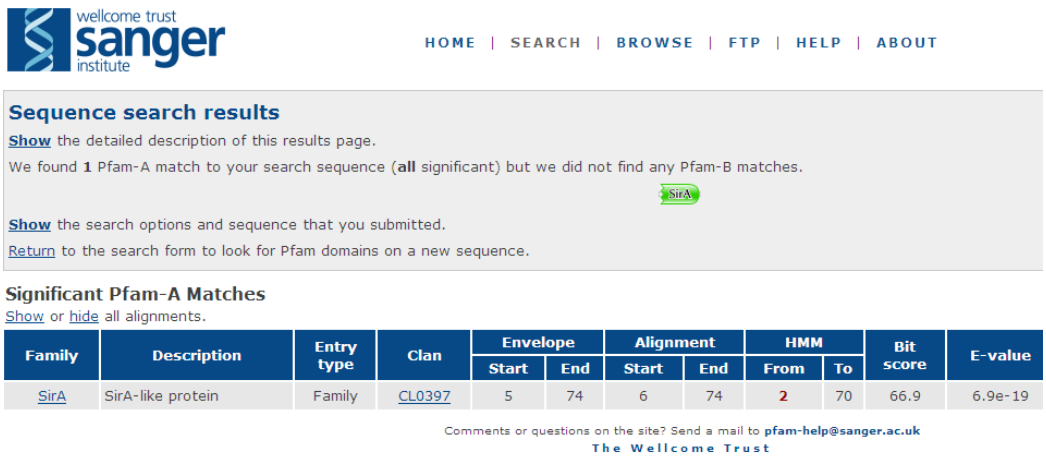
- Is there any significance between the location of Cj1500-01 and the FDH genes (Cj1509-11c)?

- What are the roles of Cj1500 and Cj1501 (based on bioinformatic data)?

9.3 Results & Discussion

9.3.1 Cj1501

Pfam is a database of protein families that includes their annotations and multiple sequence alignments generated using hidden Markov models (http://pfam.sanger.ac.uk/ [4]). The Pfam database contains information about protein domains and families. Pfam-A is the manually curated portion of the database that contains over 10,000 entries. For each entry, a protein sequence alignment and a hidden Markov model is stored. When the amino acid sequence of Cj1501 was examined in the Pfam tool, the software detected one significant match: SirA-like protein. The software predicted, based on the amino acid sequence and using a hidden Markov model, that the protein was a single complete SirA-like domain (Figure 9.1).


wellcome trust sanger institute

HOME | SEARCH | BROWSE | FTP | HELP | ABOUT

Sequence search results

Show the detailed description of this results page.

We found **1** Pfam-A match to your search sequence (**all** significant) but we did not find any Pfam-B matches.

SirA

Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		Bit score	E-value
				Start	End	Start	End	From	To		
SirA	SirA-like protein	Family	CL0397	5	74	6	74	2	70	66.9	6.9e-19

Comments or questions on the site? Send a mail to pfam-help@sanger.ac.uk

The Wellcome Trust

Figure 9.1 Pfam result of Cj1501
The Pfam tool predicted, based on the amino acid sequence and using a hidden Markov model, that the Cj1501 protein was a single SirA-like domain

The structure of SirA consists of an alpha/beta sandwich with a beta-alpha-beta-alpha-beta(2) fold, comprising a mixed four-stranded beta-sheet stacked against two alpha-helices, both of which are nearly parallel to the strands of the beta-sheet. Several uncharacterised bacterial proteins (73 to 81 amino acid residues in length) that contain a well-conserved region in their N-terminal region show structural similarity to the SirA protein, including the *E. coli* protein YedF, and other members of the UPF0033 family.

SirA domain crystal structures have been identified in TA1414 (hypothetical protein) from *Thermoplasma acidophilum*, EC005 (hypothetical protein) and YhhP (a novel protein implicated in cell division) both from *Escherichia coli*, and also in TM006 (small hypothetical protein) from *Thermotoga martitima*. To date, a defined role for the SirA domain has not been indentified, however, the literature states that SirA functions as a response regulator as part of a two-component system, where BarA is the sensor kinase. However, some species, such a *C. jejuni* NCTC 11168, do not contain a BarA homologue, so it is unlikely in these cases that the SirA domain is part of this particular two-component system. In some species, SirA/BarA homologues are required for virulence gene expression, exoenzyme and antibiotic production, motility and biofilm formation [5]. The Cj1501 lacks a recognisable DNA binding domain, and as response regulators usually contain DNA-binding output domains it is unlikely that Cj1501 is a response regulator [6].

9.3.2 Cj1500

In *C. jejuni* NCTC 11168, the Cj1501 is preceded by the Cj1500 gene, and it was noticed that the sequential location of these two genes remained similar across Campylobacter species (Figure 9.2 determined by the String database). The conservation of the location of these genes is discussed later. Genes such as this, with conserved genomic positions, are noteworthy of study as they may possess a crucial role which has been consistently preserved throughout evolution. Dandekar *et al.* [7] noted that during systematic comparison of nine bacterial and archaeal genomes, there is low-level conservation of gene-order (and operon architecture), and interestingly, a number of gene pairs were conserved. They predicted that the proteins encoded by conserved gene pairs appeared to interact physically and suggested that this observation can be used to predict functions of, and interactions between, prokaryotic gene products. A study of eukaryotic genomes suggested that the gene pairs of metazoa and fungi which were evolutionary conserved, and divergently transcribed, were much more likely to be related by function when compared to poorly conserved gene pairs. The evolutionary conservation of gene pairs was studied with respect to relative transcriptional direction, intergenic distance and functional relationship as inferred by gene ontology [8].

Cj1500 has homology to the YeeE and YedE proteins from *E. coli*. These proteins are integral membrane proteins of unknown function. Many of the proteins from this YeeE/YedE family contain two homologous regions. These regions contain several conserved glycines and an invariant cysteine which is probably an important functional residue. Cj1500 is currently annotated as a probable integral membrane protein, with a length of 402 amino acids. According to Pfam, the Cj1500 protein contains two transmembrane domains, and a domain of unknown function (DUF395). There are also ten probable transmembrane helices in Cj1500, predicted by TMHMM2.0., which is a software program used to predict transmembrane protein topology using a hidden Markov model (http://www.cbs.dtu.dk/services/TMHMM/) (Figure 9.3).

9.3.3 The Cj1500-01 homologue; C8J_1404 a 'fusion' protein

Interestingly, the Cj1500-01 genes, which occur adjacently in *C. jejuni* NCTC 11168, occur as a single fused gene in *C. jejuni* 81116 (Figure 9.2). *C. jejuni* 81116 is the only example of this fused protein. When the entire C8J_1404 amino acid sequence was subjected to pBLAST search, the possible sequences with homology identified by the tool all aligned with approximately the first 400 amino-acids (Cj1500 homologous region) of the protein only. This suggested that another protein, similar to the entirety of C8J_1404, does not exist. The pBLAST tool works in a similar way to the BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), as mentioned before, and specifically searches a protein database using a protein query.

C8J_1404 is 479 amino acids long and is annotated as a putative inner membrane protein, which is consistent with the annotation of Cj1500 and Cj1501. It is most likely a transmembrane protein (similar to Cj1500), with the Cj1501-homologous region at the C-terminal. Sequence analysis using TMHMM2.0, which predicts transmembrane helices in proteins, suggested that the C-terminal (Cj1501 homologous region) is located inside the *C. jejuni* cytoplasm. This is shown diagrammatically in Figure 9.3. This information may be useful in the prediction of a role of the C8J_1404 homologues; Cj1500 and Cj1501.

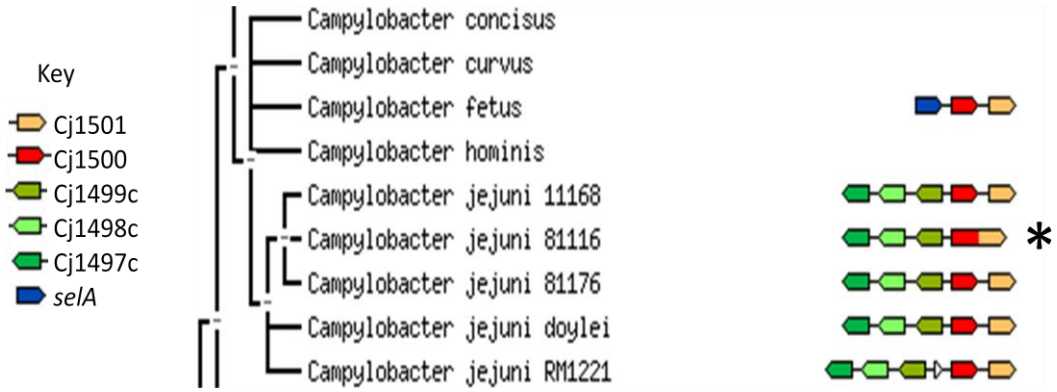


Figure 9.2 String Database analysis of Cj1500 and Cj1501
A representation of the relationship between Cj1500 (red) and Cj1501 (beige) homologues in the Campylobacter species. The image was generated by http://string-db.org/ and is based upon the location of genes spatially within the genome (not to scale). The fused genes in *C. jejuni* 81116 can clearly be identified (*). (This is the same diagram as Figure 8.4.)

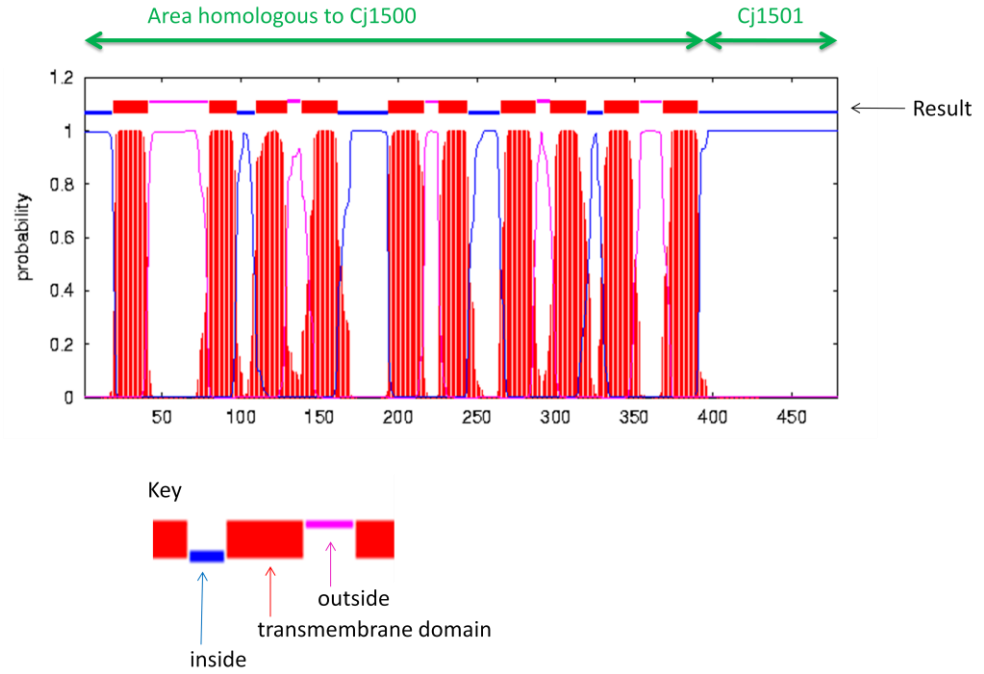


Figure 9.3 TMHMM2.0 prediction of C8J_1404
The diagrammatic prediction of the transmembrane, inner and outer domains of C8J_1404, using the TMHMM2.0 tool (http://www.cbs.dtu.dk/services/TMHMM-2.0/). The line at the top of the graph clearly depicts the amino acids from approximately 400-480 (the Cj1501 homologous region) as **blue** (within the cytoplasm).

9.3.4 Genes Involved with Selenenium Incorporation and Metabolism

Figure 9.4 diagrammatically describes the mechanism for selenocycteine formation and incorporation. In many biological systems, selenium can be incorporated in a specific pathway dedicated to the biosynthesis of proteins which contain the amino acid selenocysteine as directed by the UGA codon. In *E. coli,* this pathway requires the products of four genes, *selA*, *selB*, *selC* (the tRNA) and *selD*. The pathway is summarised in Figure 9.4. Relatively little is known about the selenocysteine formation and incorporation in *C. jejuni* NCTC 11168. Three *C. jejuni* genes have been annotated as *selA* (Cj1378), *selB* (Cj1379) and *selD* (Cj1504c). The genes were all annotated in *C. jejuni* based on sequence similarity to the *selABD* genes from *E. coli*, which have been well characterised.

The *C. jejuni selD* has a length of 308 amino acids, and a high similarity to the *selD* in *E. coli* [9]. The gene product of *selD* is a selenophosphate synthetase which catalyses the conversion of selenium to selenophosphate, the activated selenium donor, which is required by a number of bacterial, archaeal and eukaryotic organisms for synthesis of $tRNA^{Sec}$, the precursor of selenocysteine in selenoenzymes. Selenophosphate also is required as the donor form of selenium required for the conversion of 2-thiouridine to 2-selenouridine in tRNA.

The *selB* gene product, a specialised translation factor binds guanine nucleotides, selenocysteyl-tRNA, and the SECIS element, a unique mRNA stem loop structure in the mRNA located immediately downstream of UGA. The binding of the *selB* translation factor complex to the stem loop stalls the ribosome at the UGA and prevents termination of protein synthesis long enough for the insertion of the selenocysteine residue to occur. The *selA* gene product, selenocysteine synthetase, converts the $tRNA^{Sec}$ which initially is aminoacylated with serine into selenocysteyl-$tRNA^{Sec}$.

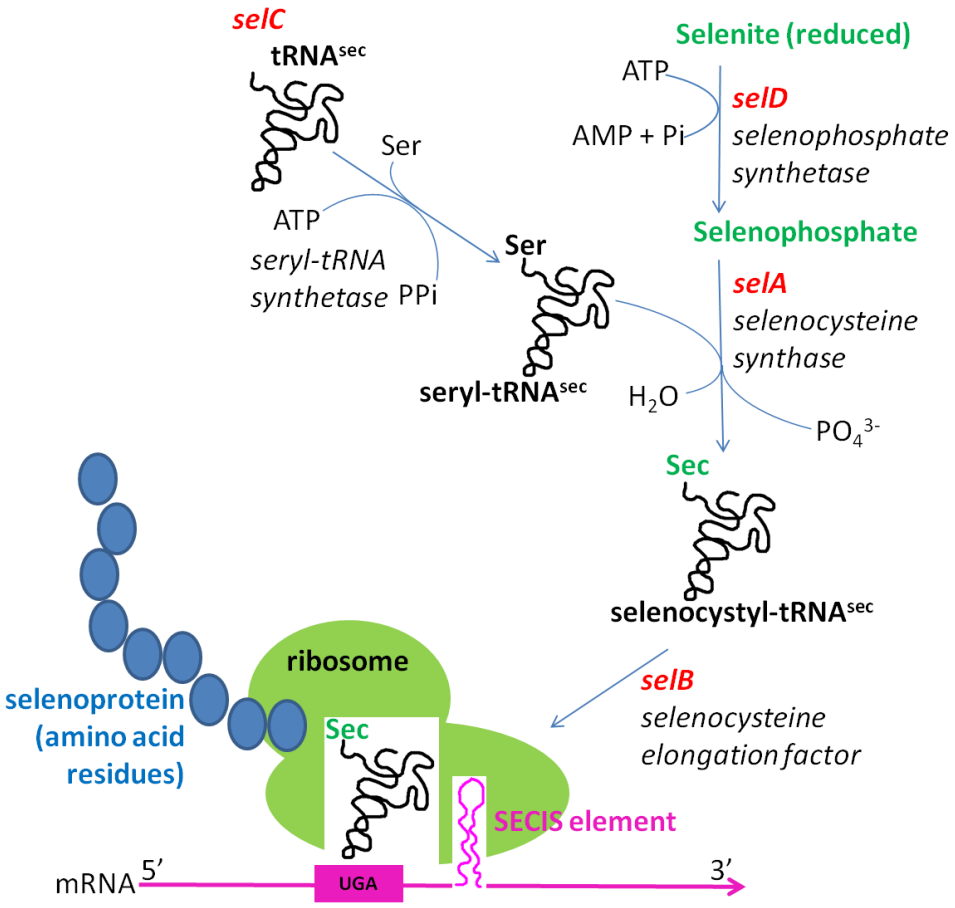


Figure 9.4 A generic diagrammatic representation of selenocysteine incorporation

In addition to *selABD*, there is a small length of *C. jejuni* NCTC 11168 genome identified as *selC*. *selC* is the term given to the $tRNA^{Sec}$ which is the special tRNA that inserts selenocysteine at certain in-frame UGA codons.

Interestingly, *selD* (Cj1504c) is located in the *C. jejuni* NCTC 11168 genome closer to the FDH genes (Cj1511-09c) and the Cj1501 gene, than *selA* and *selB* genes (Cj1378 and Cj1379 respectively). The *selC* homologue in *C. jejuni* NCTC 11168 is the 98 base pairs located upstream of the Cj1500 on the complementary strand of the genome (Figure 9.5). However, the *C. jejuni* genome organisation of gene clusters is often unclear [10], so the apparent 'closeness' of the genes in discussion here may be an artefact and should not be analysed too closely.

9.3.5 Genes Required for Formate Dehydrogenase (FDH)

Figure 9.5 summarises the genes in the FDH operon diagrammatically. The role of FDH is to catalyse the oxidation of formate to carbon dioxide, and, in *C. jejuni,* formate is a substrate which bacterial cells can metabolise for energy.

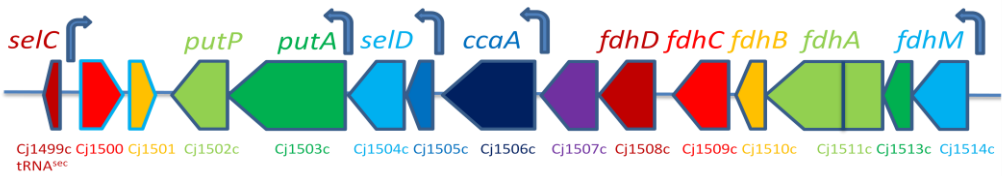


Figure 9.5 The diagrammatic representation of the FDH operon and surrounding genomic area.
The names, gene annotations and predicted promotor locations are all depicted.

Cj1511c: The largest subunit of the FDH enzyme is Cj1511c; it codes for a tungsto-selenosubunit, is known as *fdhA* and has a length of 934 amino acids. Homologues of Cj1511c (*fdhA*) from *E. coli* and *Wolinella succinogenes* have been characterised and have a high similarity to the *C. jejuni* gene. Interestingly, a UGA codes for selenocysteine at codon 181. This selenocysteine is most likely the amino acid which co-ordinates to the tungsten [11, 12]. The *fdhA* subunit has also been identified as a possible TAT (twin-arginine transporter) secreted protein. The TAT system exports folded proteins from the bacterial cell [13]. The TAT-exported proteins contain a signal motif recognised by the transporter, this is present in the Cj1511c gene (Figure 9.6). Therefore, FDH is most likely a periplasmic enzyme because it is probably exported from the cytoplasm via the TAT apparatus.

Cj1510c: Cj1510c (*fdhB*) is a smaller (213 amino acids) FDH subunit and is annotated as a probable iron-sulfur subunit. Searching the Pfam database reveals the protein has a match to entry PF00037 fer4, 4Fe-4S ferredoxins and related iron-sulfur cluster binding domains. The *fdhB* homologue in *W. succinogenes* has been characterised and has some similarity with the *C. jejuni* gene.

Cj1509c: The other subunit of the FDH protein is thought to be Cj1509c (*fdhC*), which is 310 amino acids long and annotated as a probable FDH cytochrome B subunit. The *C. jejuni* gene has similarity to *W. succinogenes fdhC* (306 amino acids), which has been charactertised. The Cj1509c contains six probable transmembrane helices predicted by TMHMM2.0, so may have a function which requires it to be tranmembrane located. Cytochromes are, in general, membrane-bound hemoproteins and carry out electron transport. Often they are found as subunits of bigger enzymatic complexes that catalyse redox reactions, such as the FDH. Cytochromes have been identified in Campylobacter with respect to different respiratory systems, including formate dehydrogenase [14]. It is a reported fact that TAT exported enzymes may also be integral membrane proteins [15]. Therefore, it is likely that after being exported through the TAT system, the FDH is anchored into the cytoplasmic membrane by the *fdhC* subunit (Cj1509c).

Cj1508c: In addition to the three FDH subunits, some proteins have been identified as required for FDH activity in micro-organisms. Cj1508c (*fdhD*) is 260 amino acid residues long and similar to the *fdhD* protein in *E. coli* (277 amino acids). In *E. coli, fdhD* is required for wild-type FDH activity, but not for synthesis of the FDH enzyme [16]. The *C. jejuni fdhD* has not been characterised, so the exact role of Cj1508c remains unknown. In *E. coli*, *fdhD* does not control transcription or UGA decoding of the FDH structural genes [17]. *E. coli* contains a *fdhE* gene, which is also required for FDH enzymatic activity, but not for transcription of the FDH operon [16]. However, there is no recognisable *fdhE* homologue in *C. jejuni* NCTC 11168.

Cj1514c: The gene Cj1514c has been shown to be specifically required for the activity of FDH and was designated *fdhM* [18]. Most periplasmic redox enzymes with complex co-factors, which FDH may be, have a dedicated redox enzyme maturation protein (REMP) that partly serve to co-ordinate co-factor insertion in the cytoplasm with TAT translocation [18]. In *C. jejuni* NCTC 11168, only two such REMPs are encoded in the genome [19], Cj1514c and Cj0785 (*NapD*). Cj1514c was predicted to be the REMP for FDH and is located upstream of Cj1511c (*fdhA*) the major subunit of FDH. In other bacteria, the FDH REMP has yet to be clearly identified, but in *C. jejuni,* REMP Cj1514c was

designated *fdhM* (for maturation). Furthermore, Cj1514c has a complex overlapping gene arrangement with a novel small gene, Cj1513c, which also overlaps *fdhA* (Cj1511c). Cj1513c is composed of a TAT signal sequence with a strongly predicted cleavage site that would give a mature peptide of just 33 residues. So, it is most likely the protein is also located in the periplasm and exported through the TAT apparatus. The role of this novel protein is unknown but it seems to be specific to the ε-Proteobacteria. One intriguing possibility is that this protein is somehow related to the likely use of a tungsten-pterin co-factor in the *C. jejuni* FDH [18].

Cj1506c: Cj1506c is annotated as probable methyl-accepting chemotaxis protein (MCP) type signal transduction protein with a length of 700 amino acids. Cj1506c contains two possible transmembrane domains and a Pfam match to entry PF00015 MCP signaling domain. The gene has been characterised [20] and is currently thought to be an aspartate receptor (*ccaA*) and belongs within the transmembrane sensory proteins, known as transducer-like proteins. However, given the location of the gene in the genome (Figure 9.5), and the fact that an aspartate molecule contains two peripheral formate-like groups (Figure 9.7), it is possible that the protein also recognises formate using the same mechanism. There is evidence to suggest that *C. jejuni* possesses chemotactic ability towards formate [21].

Cj1505c & Cj1507c: Cj1505c and Cj1507c are currently not thought to be involved with the FDH synthesis. Cj1505 is 190 amino acid residues long and annotated as a SirA-like protein (similar to Cj1501, which is a single SirA domain), but currently is of unknown function. Cj1507c is annotated as a putative regulatory protein of 245 amino acid residues. Pfam identifies Cj1507c as containing a regulatory helix-turn-helix protein domain, and suggests it may belong to the LysR family. However, a specific function has yet to be assigned to the Cj1507c protein. It is possible that Cj1505c is responsible for the transcriptional regulation of the FDH genes, but to determine this, further experiments with genetic manipulation of the Cj1505c would be necessary.

Cj1503c & Cj1502c: Cj1503c and Cj1502c have been annotated *putA* and *putP,* putative proline dehydrogenase and a putative sodium/proline symporter, respectively. There is no information in the databases to suggest that *putA* and *putP* have roles in selenium metabolism and incorporation, or FDH formation and activity.

Promoters: The locations of the promoters in this genome region, determined experimentally by differential RNA-seq, (Arnoud van Vliet, IFR, method from [22]) are shown in Figure 9.5. It is predicted that the FDH subunit genes (*fdhABC*) are transcribed at the same time as Cj1514c and Cj1513c, under a single promotor. Transcriptional start site analysis did not indicate that *fdhD* has its own promoter, however, it is likely that *fdhD* has a promoter and that it co-transcribes with Cj1507c because there is a gap in front of Cj1508c which could be a promoter. The *putA* and *putP* are transcribed together, as are Cj1505c and *selD.* The Cj1500-01 genes, which are located on the opposite strand to the other genes in the locality, are regulated under their own promoter. The occurrence of the Cj1500-01 genes in the opposite direction to all the other genes within the region suggests that perhaps the genes were acquired in *C. jejuni* NCTC 11168 by horizontal gene transfer.

Protein	Predicted function	N-terminal sequence*
→ Cj1513c	Periplasmic protein	VKN**RREFLK**KSAFALGAAGVLGSTTLALAKDEERKDLVKG
Cj0379c	Unknown	MLITPEKLYKQ**RRNFLK**LGAGALISSSVLASKLSALNFTSDT
Cj0005c	Molybdenum containing oxidoreductase	MKQNDQKEN**RRDFLK**NIGLGLFGISVLSNFSFENFLGSKALA
Cj1186c	Ubiquinol-cytochrome C reductase iron–sulfur subunit	MATSES**RRSFMG**FAFGSVAAVGGVFSLVAMKKTWDPLPSVK
Cj0414	Oxidoreductase subunit	MQDNIID**RRSFFK**LGLLGGSVVAASTIGGGAVLKAAELTHSH
Cj1516	Periplasmic oxidoreductase	MN**RRNFLK**FNALTLASMGVAYANPMHDMHMHKNHSINHDL
Cj0145	Unknown	ME**RRLFLK**GSALGSMVAFFASSNLSAAMLKDKDLLGFKAVS
Cj0780	Nitrate reductase	MN**RRDFIK**NTAIASAASVAGLSVPSSMLGAQEEDWKWDKAV
Cj0781	Ferredoxin	MKG**RREFFV**SAFKAACLCTGGGFLANLALKADDNYALRPPG
→ Cj1511c	Formate dehydrogenase large subunit	MSSVGENIKLT**RRSFLK**MAALSSLATPLLARSETLREASADEL
Cj0264c	Molybdopterin-containing oxidoreductase	MLD**RRKFLK**IGASLSALPLIPSLSAGKTVEASKVSLGLVKNGE
Cj0437	Succinate dehydrogenase flavoprotein subunit	MGEFS**RRDFLK**TACISVGALAASSSGVYALDDSSKMDKDINL

*Potential twin-arginine recognition motifs (consensus RRXFLK) are underlined and in bold. Additional TAT-dependent proteins harbouring less conserved signal sequences may also exist.

Figure 9.6 The *C. jejuni* NCTC 11168 genes encoding proteins that are potentially secreted by the TAT apparatus (taken from [13]).
Cj1511c and Cj1513c are highlighted with red arrows.

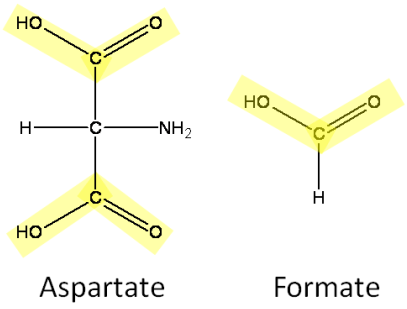


Figure 9.7 The structure of aspartate and formate, with the similar carboxylic acid (COOH) groups highlighted.

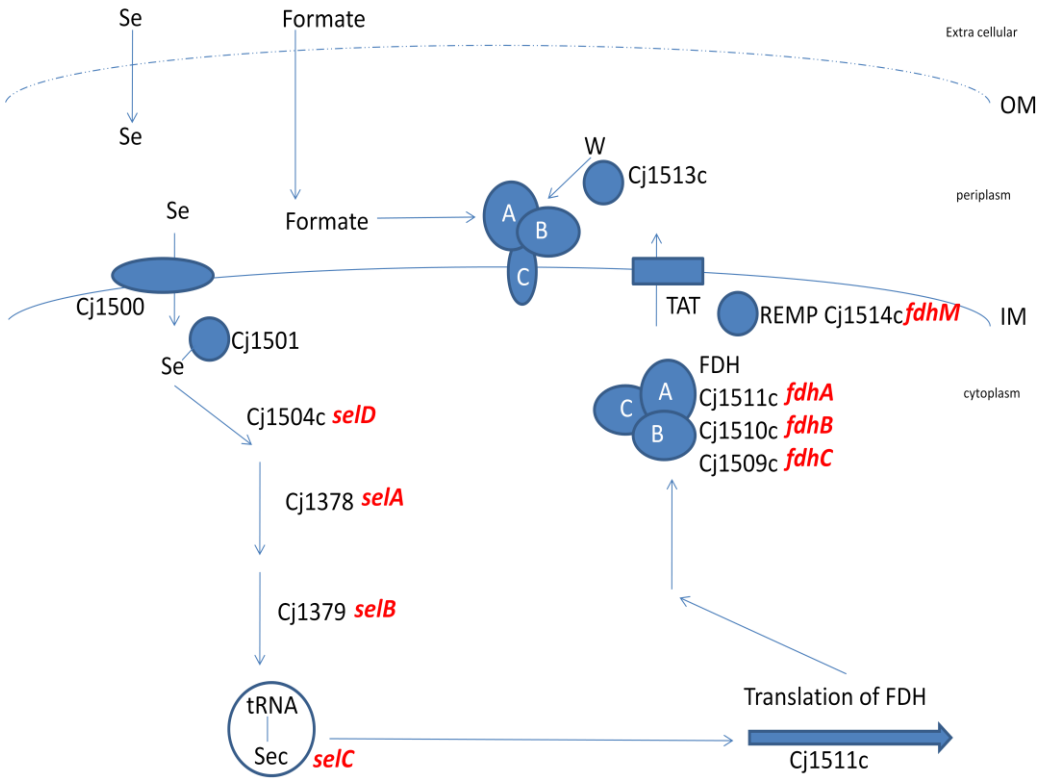


Figure 9.8 The possible function of the *C. jejuni* NCTC 11168 genes relating to the production of active FDH.
Gene names shown in red. W – tungsten, Se – selenium, OM – outer membrane, IM – inner membrane, REMP – redox enzyme maturation protein, TAT – twin-arginine transporter, FDH – formate dehydrogenase.

9.3.6 Cellular prediction of genes involved with FDH activity in *C. jejuni*

Figure 9.8 shows a schematic representation of the genes possibly required for production of a functioning FDH enzyme in *C. jejuni* NCTC 11168 and their possible cellular roles within the FDH production process. The first step in the process is the movement of both formate and selenium from the extracellular environment into the periplasm or cytoplasm, respectively. The periplasm is a region in the cell wall of Gram-negative bacteria, between the outer membrane and cytoplasmic membrane. The periplasm is recognised as an important, specialised region; hydrolytic, degradative and electron transport enzymes are often located within the periplasm, and typically it is more viscous than the cytoplasm. In addition, the active sub-unit of FDH of *W.*

succinogenes is found in the periplasm, anchored to the cytoplasmic membrane [23].

In *C. jejuni* NCTC 11168, Cj1500 is predicted to be located in the cytoplasmic cell membrane and it is hypothesised, based on the transmembrane domain prediction, that it is a transporter or membrane channel. When Cj1500 was inactivated (Chapter Eight), the cell did not synthesise a detectable active FDH protein (analysed by the benzyl viologen linked assay). However, FDH activity was restored by the addition of exogenous selenium. Therefore, it is possible that Cj1500 is a selenium transporter and, under normal conditions, it moves selenium into the cytoplasm. Selenium is required for a fully functioning FDH, which is a selenoprotein.

When the selenium is located intracellularly, Cj1501 may have a role in moving the selenium to the selenocysteine synthase apparatus. Cj1501 is most likely a cytoplasmic enzyme based on the sequence similarity to the transmembrane homologue (C8J_1404) from *C. jejuni* 81116. It is presently unknown whether Cj1501 is able to bind selenium either as an ion or in another form.

Cj1504c (*selD*), Cj1378 (*selA*) and Cj1379 (*selB*) then change the selenium ion into a selenocysteine, which is the form required for protein incorporation, as previously discussed. The selenocysteine is inserted into the protein sequence during the translational process of Cj1511c. The three sub units Cj1511c (*fdhA*), Cj1510c (*fdhB*) and Cj1509c (*fdhC*) make up the FDH enzyme. The FDH is exported from the cell via a TAT transporter and assisted by REMP Cj1514c (*fdhM*). It is then possible that Cj1513c, which is also TAT exported to the periplasm, assists in the insertion of the tungsten-pterin co-factor which FDH contains. It is most likely that the dehydrogenation of formate takes place in the periplasm, with the FDH anchored onto the cytoplasmic membrane by the Cj1509c subunit.

9.3.7 Selenium Binding Proteins (SBPs) and Selenium Binding Motifs

The intracellular movement of selenium before incorporation into a selenoprotein is discussed here. Selenophosphate, an activated form of selenium that can serve as a selenium donor, is generated by the *selD* gene product, selenophosphate synthetase. Selenophosphate is required by several bacteria and by mammals for the specific synthesis of tRNA(Sec), the precursor of selenocysteine in selenoenzymes. Although free selenide can be used *in vitro* for synthesis of selenophosphate, the physiological system that donates selenium to *selD* protein is incompletely characterised [24]. Using bioinformatic tools, the possibility of Cj1501 acting in this uncharacterised biological system is explored.

Small selenium-binding proteins do exist; a 14 KDa selenium-binding protein was characterised from mouse liver [25]. There is a clear difference between the selenoproteins which contain selenium in the form of a selenocysteine, such as *C. jejuni* FDH, and selenium-binding proteins (SBPs), which tightly bind a selenium ion [24]. Two known sulfurtransferases, Rhodanese (EC 2.8.1.1) and 3-Mercaptopyruvate sulfurtransferase (EC 2.8.1.2), and a glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) have been analysed for ability to bind and transfer selenium. The results indicated that all three proteins were able to tightly bind selenium [26]. An investigation in *E. coli* revealed the presence of five selenium binding proteins (Figure 9.9). Interestingly, a 7.3 KDa protein (which was similar size to Cj1501, 75 amino acid residues, 8 KDa) was identified as selenium-binding. The 7.3 KDa protein (67 amino acids) was identified as YDFZ hypothetical protein and exhibited homology to a nitrogenase protein from *Clostridium pasteurianum* [27]. However, the *E. coli* YDFZ and *C. jejuni* Cj1501 protein sequences have very limited similarity to one another when compared using the SIM program (http://expasy.org/tools/sim-prot.html [28]), which finds a user-defined number of best non-intersecting alignments between two protein sequences or within a sequence (Appendix V).

Other proteins which have also been identified as selenium-binding include cysteine desulfurase (CsdB), cysteine sulfinate desulfinase (CSD) and iron

sulfur cluster (IscS) proteins function both *in vitro* and *in vivo* as selenium delivery proteins in the wild-type *E. coli* and also in a *selD E. coli* mutant [29]. The three cysteine desulfurase-like proteins (CsdB, CSD and IscS) catalysed the removal of sulfur and selenium from L-cysteine and L-selenocysteine, to form L-alanine. These enzymes are proposed to function as sulfur-delivery proteins for iron-sulfur cluster, thiamine, 4-thiouridine, biotin and molybdopterin. It was reported that selenium mobilised from free selenocysteine was incorporated specifically into a selenoprotein and tRNA *in vivo*, supporting the involvement of the cysteine desulfurase-like proteins in selenium metabolism.

Mass (kDa)	Identified protein
90	ND
60	ND
39	GAPDH*
28	Deoxyribose-5-phosphate aldolase (DPA)‡
7.3	YDFZ hypothetical protein‡

Figure 9.9 A table depicting the identities of *E. coli* proteins with the ability to bind selenium.
Taken from Lacourciere's investigation regarding the detection of potential selenium delivery proteins using an *E. coli* strain unable to incorporate selenium from selenite into proteins [29].

Plant or mammal SBP homologues have been reported as containing a bis(cysteinyl) sequence motif, Cys-X-X-Cys [30]. It is believed that this motif is where the selenium ion interacts with the SBP. The conserved Cys-X-X-Cys motif is present in both plant and animal SBP proteins and contains two adjacent cysteines separated by two amino acid residues; this is also a characteristic feature of several proteins that are involved in controlling the oxidation/reduction status (redox proteins) of target proteins *in vivo* [31]. The Cys-X-X-Cys motif is not present in Cj1501 protein sequence, or interestingly the *E. coli* YDFZ protein described as selenium binding in *E. coli* (Figures 9.9 & Appendix V).

The amino acid sequences of all the putative SBPs compared by Flemetakis *et al.* [32] contained a high percentage (about 3%) of histidine residues, some of which appear to be part of putative metal-binding motifs of the His-X-Asp or

the His-X-X-His types. However, there are no histidine residues in the Cj1501 protein sequence. Therefore, if Cj1501 and YDFZ are SBPs, there must be an alternative, previously uncharacterised, selenium binding site used by these proteins. This indicates that the mechanism, function and role that Cj1501 is performing intracellularly is definitely a research-worthy subject. The Cj1501 clearly has a role in FDH formation, and perhaps is involved with selenium metabolism or incorporation, but has no similarity to proteins previously reported in these roles in the literature.

It is difficult using the literature alone to determine whether Cj1501 is a selenium binding protein. It is known that proteins of this size are able to bind selenium, and that selenium binding proteins are essential in the movement of selenium intracellularly, and the removal of these can result in inactive selenoproteins, such as *C. jejuni* FDH. However, the research regarding selenium-binding motifs and domains is not currently appropriate to the Cj1501 protein. It is possible that the Cj1501 does bind selenium, but using a currently uncharacterised method. The only way to undoubtedly confirm whether Cj1501 binds selenium, would be the heterologous expression and purification of the protein, followed by selenium-binding assays. Always, it is more robust to have the function of a gene characterised biochemically, then no doubt exists as to whether the bioinformatic analysis has been performed correctly, or if the links made have been established based on inadequate information.

9.3.8 Investigation of phylogenetically diverse organisms which contain both Cj1500 and Cj1501 adjacent homologues and a selenoprotein, such as FDH.

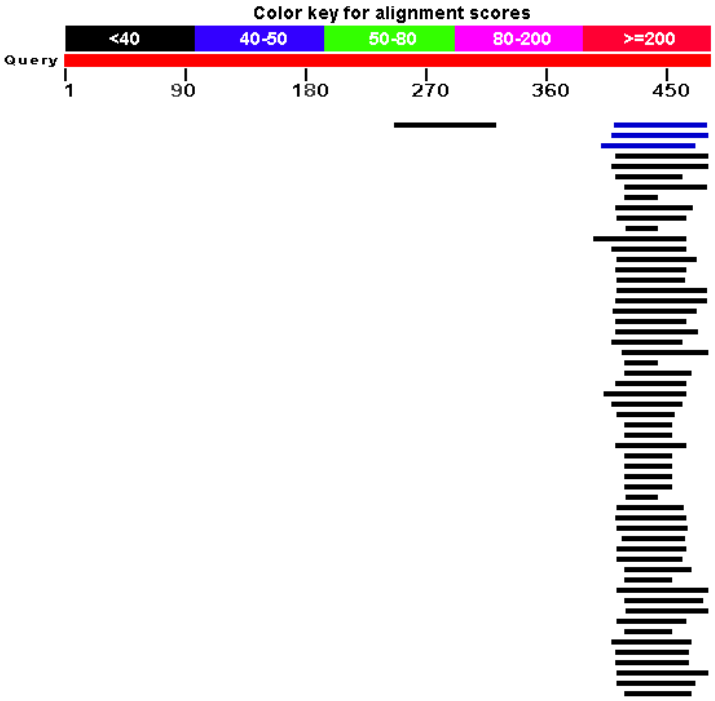
To investigate whether Cj1500 and Cj1501 have a role in selenium metabolism and incorporation into the *C. jejuni* FDH, it would be interesting to discover if the Cj1501, or the Cj1500-01 complex, is conserved across those species which contain a selenoprotein.

When the protein sequence of Cj1500-01 was subjected to a pBLAST, the sequences which had the closest similarity were mainly annotated as

'hypothetical' proteins. This information is not particularly useful when trying to determine the role of a protein bioinformatically. However, when Cj1500-01 was subjected to a pBLAST search and the results were limited to those containing the word 'selenium,' either in the name or protein description, the results were interesting; it appeared that only the Cj1501 region of the genes had homology to selenium-related proteins (Figure 9.10). Consequently, the Cj1501 protein sequence alone was pBLAST and once again the results were limited to only selenium-containing annotations, this can be seen in Figure 9.11. The Cj1501 protein had sequence homology with a number of SirA-family proteins also. This was interesting as it agreed with the earlier domain searching using the Pfam tool, which suggested that Cj1501 was a single SirA domain.

Further investigation into the species *Clostridium sticklandii, Clostridium ljungdahlii, Fusobacterium varium, Nautilia profundicola, Eubacterium hallii, Desulfotomaculum acetoxidans, Ruminococcaceae bacterium* and *Fusobacterium ulcerans* which all contained a Cj1501-like protein homologue, annotated with the word selenium, revealed that the relationship between Cj1501 and selenoprotein to be inconsistent. Table 9.1 shows the findings of this investigation. The bioinformatic analysis highlighted the presence of bacteria similar to *C. jejuni*, such as *C. sticklandii* and *C. ljungdahlii,* which contained both a Cj1501 homologue (selenium metabolism protein, YedF) and a large selenocysteine containing FDH subunit. It is possible that Cj1501 is a selenium metabolism protein (like those of *C. sticklandii* and *C.ljungdahlii*) and is performing the same role in *C. jejuni* as the Clostridia species. The selenium metabolism protein, YedF, from the Clostridia species, has not been biochemically profiled and is annotated based on the method of partial phylogenetic profiling. This YedF protein family is found only among those genomes that also carry the selenium donor protein *selD*, and its connection to selenium metabolism is indicated only by the method of partial phylogenetic profiling. In most organisms, the YedF gene is typically found next to *selD*. In the *C. jejuni* NCTC 11168 genome, Cj1501 is two genes away from *selD* (Cj1504c), however, it is noteworthy to point out that the genes are found on opposite strands of the chromosome (Figure 9.5), unlike the previously described gene pair (*yedF* and *selD*).

The role of YedF in selenium metabolism is unclear, but may include either detoxification or a role in selenoprotein biosynthesis. Members of this YedF family are found even when selenocysteine and selenouridine biosynthesis pathways are, except for *selD*, completely absent, as in *Enterococcus faecalis*. This is a noteworthy point; this perhaps explains why homologues of Cj1501 are found in cells which appear not to have a selenoprotein. For example, *N. profundicola* and *E. hallii* both have Cj1501 homologues, and the genes required for selenocysteine formation, however their FDH is not selenocysteine containing, and it is currently unknown whether the cells contain an alternative selenoprotein. If the organisms do contain an alternative selenoprotein, it is possible that the Cj1501-like proteins are able to aid selenocysteine incorporation into other selenoproteins, and not limited to FDHs alone. When the genome annotation is relatively new, and little investigation has been performed on the organisms in question, it is difficult to determine whether or not a selenoprotein is present. Software can scan sequenced genomes are predict the possibility of selenoproteins occurring by identifying the SECIS structure downstream of the UGA which co-ordinates the insertion of a selenocysteine. However, if an alternative selenoprotein is not within the cell, this raises interesting questions regarding the function of the Cj1501-like protein, such as is the small protein binding selenium for another use, or binding an alternative metal ion? Perhaps the Cj1501-like protein does have a role in regulation, as the initial Pfam search suggested when it was discovered to be a single SirA domain, which are consistently found in regulatory proteins. To definitively answer these hypotheses, biochemical testing would be necessary in addition to bioinformatic analysis of sequenced and annotated genomes.

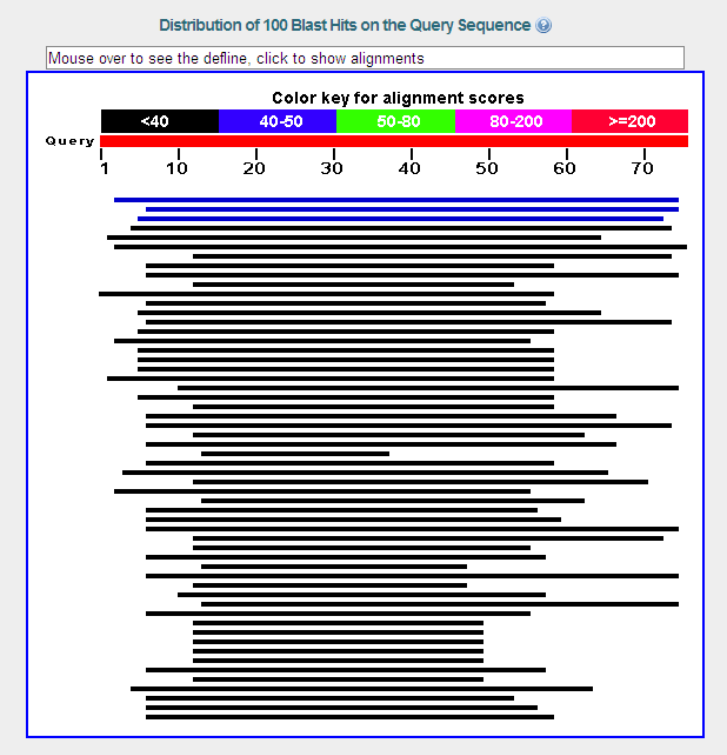


Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Links
ZP_01965120.1	hypothetical protein RUMOBE_02851 [Ruminococcus obeum ATCC 29	41.6	41.6	14%	1e-04	
ZP_07928010.1	conserved hypothetical protein [Fusobacterium ulcerans ATCC 4918	40.4	40.4	15%	3e-04	
ZP_06346510.1	selenium metabolism protein YedF [Clostridium sp. M62/1]	40.0	40.0	14%	4e-04	
ZP_05633372.1	hypothetical protein FulcA4_08042 [Fusobacterium ulcerans ATCC 4	39.7	39.7	14%	4e-04	
ZP_04858769.1	conserved hypothetical protein [Fusobacterium varium ATCC 27725]	38.5	38.5	15%	0.001	
ZP_07634553.1	SirA family protein [Ruminococcaceae bacterium D16]	37.4	37.4	10%	0.002	
YP_002607685.1	hypothetical protein NAMH_1292 [Nautilia profundicola AmH]	37.0	37.0	12%	0.003	G
ZP_03568714.1	selenium metabolism protein YedF [Atopobium rimae ATCC 49626]	37.0	37.0	5%	0.003	
YP_001952682.1	SirA family protein [Geobacter lovleyi SZ]	37.0	37.0	12%	0.003	G
ZP_03715712.1	hypothetical protein EUBHAL_00769 [Eubacterium hallii DSM 3353]	36.6	36.6	10%	0.004	
YP_003179271.1	SirA family protein [Atopobium parvulum DSM 20469]	36.2	36.2	5%	0.006	G
ZP_04743822.1	selenium metabolism protein YedF [Roseburia intestinalis L1-82]	36.2	36.2	14%	0.006	
ZP_03798486.1	hypothetical protein COPCOM_00740 [Coprococcus comes ATCC 277	35.8	35.8	11%	0.006	
YP_003192452.1	SirA family protein [Desulfotomaculum acetoxidans DSM 771]	35.8	35.8	12%	0.007	G
ZP_03780515.1	hypothetical protein CLOHYLEM_07617 [Clostridium hylemonae DSM	35.4	35.4	11%	0.008	
ZP_07905393.1	selenium metabolism protein YedF [Eubacterium saburreum DSM 398	35.4	35.4	10%	0.008	
YP_003305101.1	selenium metabolism protein YedF [Sulfurospirillum deleyianum DSM 6	35.4	35.4	13%	0.009	G
ZP_01871239.1	hypothetical protein CMTB2_03613 [Caminibacter mediatlanticus TB-	35.4	35.4	14%	0.010	
YP_001212054.1	redox protein [Pelotomaculum thermopropionicum SI]	35.0	35.0	12%	0.010	G
ZP_01996157.1	hypothetical protein DORLON_02163 [Dorea longicatena DSM 13814]	34.7	34.7	11%	0.014	
YP_001112601.1	SirA family protein [Desulfotomaculum reducens MI-1]	34.7	34.7	12%	0.015	G
ZP_04856337.1	conserved hypothetical protein [Ruminococcus sp. 5_1_39B_FAA]	34.3	34.3	11%	0.021	
ZP_04567395.1	SirA family protein [Fusobacterium mortiferum ATCC 9817]	33.9	33.9	13%	0.022	
ZP_06371401.1	hypothetical protein C414_000020051 [Campylobacter jejuni subsp.	33.9	33.9	5%	0.023	
YP_384050.1	hypothetical protein Gmet_1085 [Geobacter metallireducens GS-15]	33.9	33.9	10%	0.024	G
ZP_02234498.1	hypothetical protein DORFOR_01369 [Dorea formicigenerans ATCC 27	33.9	33.9	11%	0.025	
ZP_03290473.1	hypothetical protein CLONEX_02689 [Clostridium nexile DSM 1787]	33.9	33.9	12%	0.025	
ZP_01968397.1	hypothetical protein RUMTOR_01967 [Ruminococcus torques ATCC 2	33.9	33.9	11%	0.026	
YP_891541.1	hypothetical protein CFF8240_0339 [Campylobacter fetus subsp. fet	33.5	33.5	8%	0.030	G
ZP_06374354.1	hypothetical protein C1336_000320043 [Campylobacter jejuni subsp.	33.5	33.5	7%	0.030	
YP_001482983.1	hypothetical protein C8J_1408 [Campylobacter jejuni subsp. jejuni 8	33.5	33.5	7%	0.030	G
ZP_02432011.1	hypothetical protein CLOSCI_02247 [Clostridium scindens ATCC 3570	33.5	33.5	11%	0.031	
ZP_03223414.1	hypothetical protein Cj8421_1558 [Campylobacter jejuni subsp. jejur	33.5	33.5	7%	0.033	
YP_179659.1	hypothetical protein CJE1678 [Campylobacter jejuni RM1221]	33.5	33.5	7%	0.034	G
ZP_01067536.1	conserved hypothetical protein [Campylobacter jejuni subsp. jejuni C	33.5	33.5	7%	0.034	
YP_001001152.1	hypothetical protein CJJ81176_1497 [Campylobacter jejuni subsp. je	33.5	33.5	7%	0.035	G
ZP_00371210.1	predicted Transcriptional regulator [Campylobacter upsaliensis RM31	33.1	33.1	5%	0.039	

Figure 9.10 Shows the pBLAST result of Cj1500-01 after limiting the results to only contain those with the word selenium either in the name or protein description.

The Cj1501 region is homologous with other selenium-related proteins. The 1500 part is not related.



Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Links
ZP_07928010.1	conserved hypothetical protein [Fusobacterium ulcerans ATCC 4918	40.8	40.8	96%	2e-05	
ZP_05633372.1	hypothetical protein FulcA4_08042 [Fusobacterium ulcerans ATCC 4	40.0	40.0	90%	3e-05	
ZP_07634553.1	SirA family protein [Ruminococcaceae bacterium D16]	40.0	40.0	89%	3e-05	
ZP_01965120.1	hypothetical protein RUMOBE_02851 [Ruminococcus obeum ATCC 29	39.7	39.7	92%	4e-05	
ZP_06346510.1	selenium metabolism protein YedF [Clostridium sp. M62/1]	39.7	39.7	84%	4e-05	
ZP_04858769.1	conserved hypothetical protein [Fusobacterium varium ATCC 27725]	38.9	38.9	97%	6e-05	
YP_002607685.1	hypothetical protein NAMH_1292 [Nautilia profundicola AmH]	36.6	36.6	81%	3e-04	G
ZP_03715712.1	hypothetical protein EUBHAL_00769 [Eubacterium hallii DSM 3353]	36.6	36.6	69%	4e-04	
YP_003192452.1	SirA family protein [Desulfotomaculum acetoxidans DSM 771]	36.2	36.2	90%	4e-04	G
ZP_03568714.1	selenium metabolism protein YedF [Atopobium rimae ATCC 49626]	36.2	36.2	54%	5e-04	
ZP_03798486.1	hypothetical protein COPCOM_00740 [Coprococcus comes ATCC 277	35.8	35.8	77%	5e-04	
ZP_07905393.1	selenium metabolism protein YedF [Eubacterium saburreum DSM 398	35.8	35.8	68%	5e-04	
YP_001952682.1	SirA family protein [Geobacter lovleyi SZ]	35.8	35.8	78%	5e-04	G
YP_003305101.1	selenium metabolism protein YedF [Sulfurospirillum deleyianum DSM 6	35.0	35.0	89%	0.001	G
ZP_01996157.1	hypothetical protein DORLON_02163 [Dorea longicatena DSM 13814]	35.0	35.0	70%	0.001	
ZP_04856337.1	conserved hypothetical protein [Ruminococcus sp. 5_1_39B_FAA]	35.0	35.0	70%	0.001	
ZP_03780515.1	hypothetical protein CLOHYLEM_07617 [Clostridium hylemonae DSM	34.7	34.7	70%	0.001	
ZP_04743822.1	selenium metabolism protein YedF [Roseburia intestinalis L1-82]	34.7	34.7	70%	0.001	
ZP_02234498.1	hypothetical protein DORFOR_01369 [Dorea formicigenerans ATCC 27	34.3	34.3	70%	0.001	
ZP_03290473.1	hypothetical protein CLONEX_02689 [Clostridium nexile DSM 1787]	34.3	34.3	76%	0.002	
ZP_04567395.1	SirA family protein [Fusobacterium mortiferum ATCC 9817]	33.9	33.9	85%	0.002	
ZP_02432011.1	hypothetical protein CLOSCI_02247 [Clostridium scindens ATCC 357	33.9	33.9	70%	0.002	
NP_953567.1	hypothetical protein GSU2521 [Geobacter sulfurreducens PCA]	33.9	33.9	61%	0.002	G
YP_001112601.1	SirA family protein [Desulfotomaculum reducens MI-1]	33.5	33.5	80%	0.002	G
ZP_01871239.1	hypothetical protein CMTB2_03613 [Caminibacter mediatlanticus TB-	33.5	33.5	89%	0.003	
YP_384050.1	hypothetical protein Gmet_1085 [Geobacter metallireducens GS-15]	33.5	33.5	66%	0.003	G
YP_891541.1	hypothetical protein CFF8240_0339 [Campylobacter fetus subsp. fet	33.5	33.5	80%	0.003	G
YP_003179271.1	SirA family protein [Atopobium parvulum DSM 20469]	33.5	33.5	32%	0.003	G
ZP_02036847.1	hypothetical protein BACCAP_02458 [Bacteroides capillosus ATCC 29	33.5	33.5	69%	0.003	
YP_001212054.1	redox protein [Pelotomaculum thermopropionicum SI]	33.1	33.1	82%	0.003	G
ZP_05364162.1	selenium metabolism protein YedF [Campylobacter showae RM3277]	33.1	33.1	77%	0.003	
ZP_01968397.1	hypothetical protein RUMTOR_01967 [Ruminococcus torques ATCC 2	33.1	33.1	70%	0.004	
YP_001407289.1	hypothetical protein CHAB381_1779 [Campylobacter hominis ATCC B	32.7	32.7	65%	0.005	G
ZP_03681545.1	hypothetical protein CATMIT_00157 [Catenibacterium mitsuokai DSM	32.3	32.3	66%	0.005	
ZP_02088698.1	hypothetical protein CLOBOL_06254 [Clostridium bolteae ATCC BAA-	32.3	32.3	70%	0.007	
YP_004050878.1	selenium metabolism protein yedf [Calditerrivibrio nitroreducens DSM	32.0	32.0	90%	0.009	G
NP_907203.1	hypothetical protein WS1001 [Wolinella succinogenes DSM 1740]	31.2	31.2	80%	0.012	G

score 30

Figure 9.11 Shows the pBLAST result of Cj1501 after limiting the results to only contain those with the word selenium either in the name or protein description.
The Cj1501 region is homologous with other selenium-related and SirA family proteins.

Organism name	Organism information	Cj1501-like protein annotation	FDH or selenoprotein containing
Clostridium sticklandii	Gram-positive bacterium with low (G+C) content. Anaerobic, often used as a model organism for novel biochemical reactions.	selenium metabolism protein, Yed-F (ZP_0634659.1)	A FDHα subunit (YP_003935859.1) and the recent discovery of an as yet unrecognized selenoprotein in the D-proline reductase operon [33].
Clostridium ljungdahlii	This microorganism will ingest carbon monoxide, carbon dioxide and hydrogen to produce ethanol and water.	selenium metabolism protein, Yed-F (ZP_0634659.1)	A FDHα subunit (YP_003780168.1).
Fusobacterium varium	Human pathogen which causes bacteremia.	conserved hypothetical protein (ZP_04858769.1)	Does not contain an annotated selenocysteine containing protein, however, does have probable selenium dependent proteins (EES62315.1) and selenocysteine formation genes (e.g. EES64817.1). So it is likely that it contains a seleoprotein which is yet to be identified. The FDH-like protein which has been annotated is not a selenoprotein.
Nautilia profundicola	A thermophilic, strict anaerobic, sulfur-reducing ε-proteobacterium isolated from deep-sea hydrothermal vents.	hypothetical protein (YP_002607685.1)	The FDHα subunit is not selenocysteine containing (ACM93609.1), however the genome has selenocysteine formation genes (e.g. YP_002607680.1).
Eubacterium hallii	Lactate-utilising, butyrate-producing nonsporulating gram-positive anaerobic rod-shaped bacteria normally found in soil and water. The organisms are also found in the skin and cavities of humans and other mammals, where they may cause soft-tissue infections.	selenium metabolism protein, Yed-F (ZP_03715712.1)	Contains a possible FDH, but it is not a selenoprotein. Has annotated genes involved with selenocysteine formation (e.g. ZP_03715711.1).
Desulfotomaculum acetoxidans	obligate anaerobic, acetate-oxidizing, sulfate-reducing Gram-positive sulfate-reducing bacteria known to grow with acetate as sole energy and carbon source.	SirA-family protein (YP_003192452.1)	Has two genes annotated as FDHα subunit; ACV64222.1 contains a selenocysteine, but the other gene does not.
Fusobacterium ulcerans	Obligatory anaerobic Gram-negative, non-spore-forming rod [1] that has been isolated from tropical ulcers.	conserved hypothetical protein. (ZP_0792809.1)	Does have formate metabolism, but not a FDH homologue. selenocysteine formation genes (e.g. ZP_07928009.1) So it is likely that it contains a selenoprotein which is yet to be identified.
Ruminococcaceae bacterium	Mesophilic, unclassified Ruminococcaceae.	SirA-family protein (ZP_07634553.1)	Does contain selenium dependent protein (ZP_07634550.1) and selenocysteine synthase proteins (ZP_07634549.1) Does have formate metabolism, but not a FDH homologue.

Table 9.1 The findings of the bioinformatic analysis using Cj1501-like proteins and determining the selenoprotein status of these species.
The organisms were identified by blast search and are all depicted in Figure 9.11. The quality coverage and E value of each SirA-like protein can be seen.

9.3.9 Distribution of Cj1500-01 homologues across species

To determine the sequential co-occurrence of the two genes, Cj1500 and Cj1501, across different species, the String (http://string-db.org/) database was used (Figure 9.12). This indicated that the sequential occurrence of the genes was not limited to the Campylobacter species. Members of the β-Proteobacteria and Thermatogae contain homologues of Cj1500 and Cj1501 within the same locality of the genome. This conserved distribution of Cj1500 and Cj1501 homologues argues that the two genes require one another for the function they perform.

The String tool was also used to view the relationships between Cj1501 and other genes (Figure 9.13). The main associations between Cj1501 and other genes are geographical. Interestingly, the program illustrates the geographical relationship between Cj1501 and *selA*. And in addition, the Cj1500 has both neighbourhood and co-occurrence relationships with the *selA* gene. This information is in agreement with the hypothesis that Cj1500-01 may be involved with selenium incorporation or metabolism as previously discussed.

Using *Carboxydothermus hydrogenoformans,* a member of the clostria species, and *Nitrosomonas europaea,* the β-Proteobacterium, protein sequences of the Cj1500-01 gene-pair homologues identified by the String program, an alignment was generated using the t-coffee program (http://www.ebi.ac.uk/Tools/msa/tcoffee/) between them and the Cj1500-01 protein sequence (Appendix VI). The sequences were made by taking the two protein sequences (from the homologues of Cj1500 and Cj1501) and combining them without an adjoining space. The alignment showed the protein sequences to align with each other and this confirms that the two proteins are conserved across bacterial species.

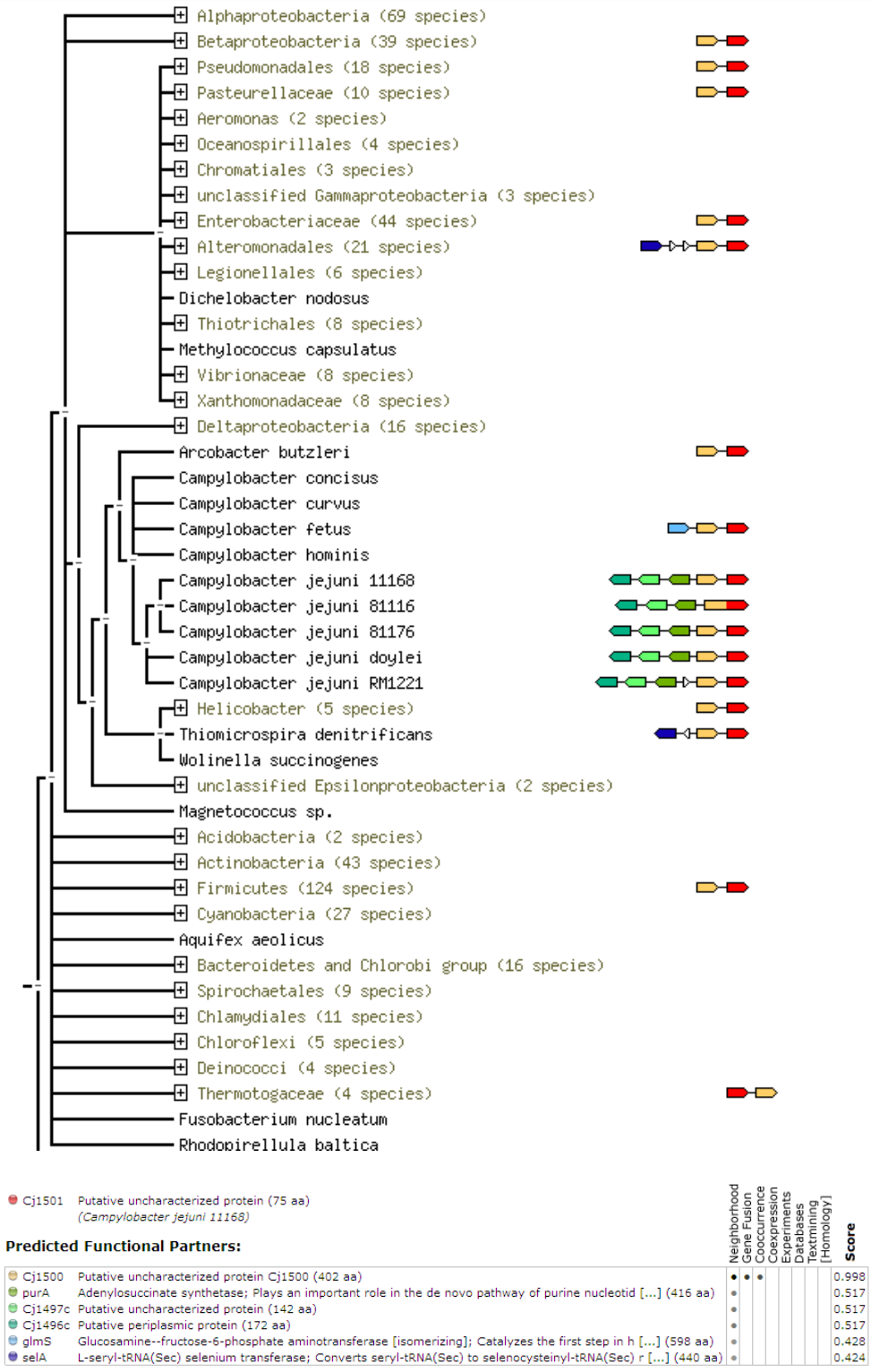
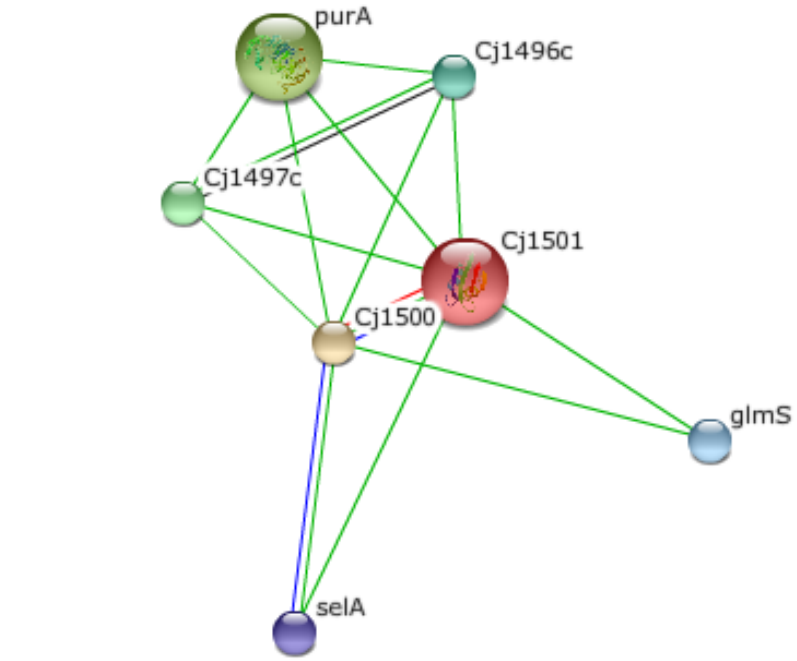


Figure 9.12 Depiction of the co-occurrence of Cj1500 and Cj1501 across different species.

The diagrammatic results are based on the occurrence of the genes in proximity to one another in the genome.



Your Input:

Cj1501	Putative uncharacterized protein (75 aa) *(Campylobacter jejuni 11168)*	

Predicted Functional Partners:

		Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
Cj1500	Putative uncharacterized protein Cj1500 (402 aa)	●	●	●						0.998
purA	Adenylosuccinate synthetase; Plays an important role in the de novo pathway of purine nucleotid [...] (416 aa)	●								0.517
Cj1497c	Putative uncharacterized protein (142 aa)	●								0.517
Cj1496c	Putative periplasmic protein (172 aa)	●								0.517
glmS	Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]; Catalyzes the first step in h [...] (598 aa)	●								0.428
selA	L-seryl-tRNA(Sec) selenium transferase; Converts seryl-tRNA(Sec) to selenocysteinyl-tRNA(Sec) r [...] (440 aa)	●								0.424

Figure 9.13 Shows a summary network depicting the relationships between Cj1501 and other genes.
The different line colors represent the types of evidence for the association (as shown in the key).

9.3.10 Roles of Cj1501

In addition to the String software, a network analysis tool which predicts protein-protein interactions was used to investigate Cj1501 (Figure 9.14). Data from large-scale protein interaction screens for humans and model eukaryotes have been invaluable for developing system-level models of biological processes. To facilitate studies of *C. jejuni*, the associates at Wayne State University and the Michigan Proteome Consortium set out to generate a proteome-wide protein interaction map for *C. jejuni* using the yeast two-hybrid system. The data generated can be used to determine protein expression profiles by mass spectroscopy based proteome mapping and to generate clone resources for *C. jejuni* researchers. Using high-throughput

yeast two-hybrid screens, 11,687 reproducible interactions were identified and the resulting interaction map included 80% of the predicted *C. jejuni* NCTC 11168 proteins and placed a large number of poorly characterised proteins into networks that provide initial clues about their functions. The map may be used to identify a number of conserved sub-networks by comparison to protein networks from *E. coli* and *Saccharomyces cerevisiae.* When completed, the *C. jejuni* protein-protein interaction map was one of the most comprehensive [34].

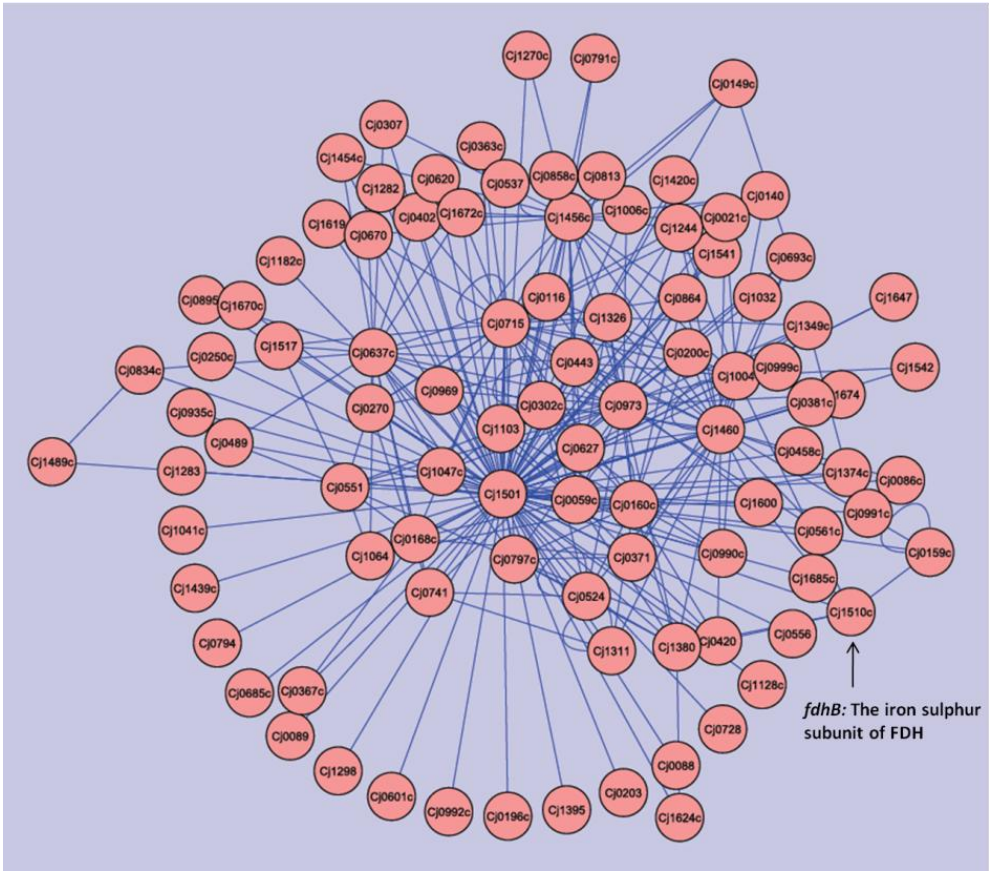


Figure 9.14 The network analysis of Cj1501.
Analysis by Mark Reuter
The diagram shows all the predicted protein-protein physical interactions. The number of interactions of Cj1501 is considered above average, this suggests that Cj1501 may be a 'hub' and have several intracellular roles.

The data generated suggests that Cj1501 is a 'hub' as it is predicted to have many interactions with different proteins (Figure 9.14). The tool is not faultless, and the data generated should only be taken as a guide, as there are bound to be false positive and negative results. The tool functions by predicting physical protein-protein interactions, and it is noteworthy to

remember that a physical interaction, or lack of, does not mean they interact with a particular pathway. The iron sulphur subunit of FDH (*fdhB*, Cj1510c) is highlighted as one of those genes which is predicted to interact with Cj1501. This is an interesting result, and if true, could suggest an interaction between Cj1501 and the FDH. Although the FDH is made of three subunits (Cj1509-11c), it would only be expected for the protein to bind in one area, therefore in one domain. A limitation of this experiment is the fact that it is based on a binary system and the results will only show interactions between two proteins. So, it cannot be predicted if Cj1501 interacts with the complete (three subunit) FDH. However, if the schematic diagram of the genes involved with FDH formation and expression (Figure 9.8) is correct, and Cj1501 is functioning as a selenium chaperone, or SBP, with a role in delivering the selenium to the selenocysteine synthesis apparatus, then it is unlikely that that Cj1501 and the FDH would interact directly. Also, it is known that FDH is a periplasmic protein (due to the TAT export motif) and Cj1501 is most likely cytoplasmic (due to the transmembrane domain structure prediction software using a Cj1501 homologue, C8J_1404), so this suggests they may not interact directly with one another, unless the interaction takes place pre-TAT transportation.

The likelihood that Cj1501 is performing as a selenium chaperone was investigated by comparison of the protein sequence with known chaperones. Using the pBLAST technique, the Cj1501 protein sequence was searched against all known sequences which contained the word 'chaperone' either in the protein name, or description. However, there were no significant hits using the software and this technique. This result suggests two possibilities; either that Cj1501 is either not a chaperone molecule, or is a completely new, uncharacterised chaperone which belongs to a family which has never been reported on, or annotated in the literature. Chaperones of 8 KDa (like Cj1501) have been reported in the literature [35]. CopZ, which is an 8 KDa protein from *Enterococcus hirae,* acts as a copper chaperone.

The Cj1501 contains a Cys-Pro-X-Pro motif towards the N terminus. This particular motif has been reported in an *E. coli* protein YhhP which is 81 amino acids in length (similar to Cj1501, which is 75). The purpose of the Cys-

Pro-X-Pro motif is to fold into a two-layered alpha/beta-sandwich structure with a beta-alpha-beta-alpha-beta-beta fold, comprising a mixed four-stranded beta-sheet stacked against two alpha-helices, both of which are nearly parallel to the strands of the beta-sheet. The Cys-Pro-X-Pro motif plays a significant structural role in stabilising the first helix. The structure of YhhP displays a striking resemblance to the C-terminal ribosome-binding domain of translation initiation factor IF3 and has an important role in cell division [36]. A variety of micro-organisms have similar proteins, all of which contain a common Cys-Pro-X-Pro sequence motif in the N-terminal region.

Interestingly, the online software tool which searches for motifs in proteins (http://myhits.isb-sib.ch/ [37]) did not highlight the Cys-Pro-X-Pro in Cj1501, nor did it predict any other metal-binding motifs. A phosphorylation motif (Thr-X-Arg) was detected at amino acids 67-69. However, it may or may not be a true phosphorylation motif; there are several known phosphorylation motifs of different lengths of amino acid residues, the shorter the motif, the higher the likelihood that it is not a conceivable site. The heterologous recombination, expression and purification of the Cj1501 protein, followed by experimentation, could confirm the presence of the phosphorylation motif.

9.4 Concluding Summary

Cj1501 has no sequence homology to previously characterised selenium binding proteins (SBPs) or known selenium binding motifs: Bioinformatic investigation of the Cj1501 concluded that it was absent of motifs similar to those observed in previously characterised SBPs. Further to this, there were no known metal binding motifs identified within the protein sequence. However, Cj1501 did contain the Cys-Pro-X-Pro motif, which has been reported as holding a significant structural role in stabilising alpha helices, and also the phosphorylation motif (Thr-X-Arg). So it possible that the Cj1501 protein has a role regarding protein structure stabilisation and may be phosphorylated.

One of two possible conclusions can be formed following the absence of a SBP motif; it can be concluded that either Cj1501 is not a SBP, or conversely, that it

does bind selenium, but in a way which is presently not fully understood. If it is an SBP it is likely that it belongs to a family of SBPs which have yet to be characterised. Either way, the specific, biochemical role of Cj1501 has yet to be proved. If, after future studies, it can be concluded that Cj1501 is a SBP, this would be the first example of a SBP of this type.

Cj1501 has no sequence homology to previously characterised chaperone proteins: Bioinformatic investigation of the Cj1501 indicated that it had no sequence homology to previously characterised chaperones. Chaperones are typically small proteins which assist in the non-covalent folding or unfolding of a protein. Regarding the lack of a functional FDH in *C. jejuni* in the absence of Cj1501, this is a plausible role for Cj1501. However, the lack of homology suggests that either Cj1501 is not a chaperone, or it is a chaperone which possesses no similarity to previously characterised chaperones. It would be interesting to determine if Cj1501 functions as a chaperone and define a specific role for the small gene, but this would call for more biochemical experiments.

The distribution of Cj1501 is not limited to those organisms containing a selenocysteine-containing FDH: Using bioinformatic tools, such as BLAST and String, it was concluded that some species which contained a Cj1501-like protein, did not possess a selenocysteine-containing FDH. It is unknown whether these species have other selenoproteins, which are presently unidentified (based on the current literature available). However, software is available to search entire genomes for the possibility of a selenoprotein. The software analyses the UGA codon and the possibility of a downstream SECIS element to determine whether a selenoprotein is likely to be formed. It would be interesting to link the distribution of Cj1501 with selenoproteins. However, within the limits of the published literature and genome annotations this was not possible. If there are examples of micro-organisms with Cj1501 homologues, but without selenoproteins, this raises an interesting question in the function of the Cj1501 protein.

Another investigation to consider would be: do those organisms which contain a selenoprotein (for example a selenocysteine-containing FDH) always contain

a Cj1501 homologue? If the answer is yes, then conclusions about the need for Cj1501 to help with selenium incorporation could be made. If the answer is no, then perhaps Cj1501 is a much more role-diverse protein than considered during this investigation.

Cj1500-01 homologues are found occurring sequentially in phylogenetically diverse species: Analysis of sequenced genomes indicated that Cj1500 and Cj1501 homologues were occurring adjacently in phylogentically diverse organisms such as members of the β-Proteobacteria and Thermatogae. When the homologues of Cj1500 and Cj1501 were found within the same locality of the genome, it was concluded that the two genes required one another for the function they perform. However, without further biochemical tests to back up this assumption, it is impossible to determine more about the role of the genes in *C. jejuni* and other micro-organisms based solely on the literature available.

Chapter Nine - References

1. Weerakoon DR, Borden NJ, Goodson CM, Grimes J, Olson JW: **The role of respiratory donor enzymes in *Campylobacter jejuni* host colonization and physiology**. *Microbial Pathogenesis* 2009, **47**(1):8-15.
2. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P *et al*: **The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored**. *Nucleic Acids Res* 2011, **39**(Database issue):D561-568.
3. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *J Mol Biol* 1990, **215**(3):403-410.
4. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K *et al*: **The Pfam protein families database**. *Nucleic Acids Res*, **38**(Database issue):D211-222.
5. Teplitski M, Goodier RI, Ahmer BM: **Pathways leading from BarA/SirA to motility and virulence gene expression in Salmonella**. *Journal of Bacteriology* 2003, **185**(24):7257-7265.
6. Galperin MY: **Diversity of structure and function of response regulator output domains**. *Curr Opin Microbiol* 2010, **13**(2):150-159.
7. Dandekar T, Snel B, Huynen M, Bork P: **Conservation of gene order: a fingerprint of proteins that physically interact**. *Trends Biochem Sci* 1998, **23**(9):324-328.
8. Davila Lopez M, Martinez Guerra JJ, Samuelsson T: **Analysis of gene order conservation in eukaryotes identifies transcriptionally and functionally linked genes**. *PLoS One* 2010, **5**(5):e10654.
9. Berg BL, Baron C, Stewart V: **Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. II. Evidence that a mRNA stem-loop structure is essential for decoding opal (UGA) as selenocysteine**. *Journal of Biological Chemistry* 1991, **266**(33):22386-22391.
10. Jagannathan A, Penn C: **Motility**. In: *Campylobacter Molecular and Cellular Biology.* Edited by Ketley J, Konkel M. Wymonham: Horizon Bioscience; 2005: 331.

11. Stolz JF, Basu P, Santini JM, Oremland RS: **Arsenic and selenium in microbial metabolism**. *Annual Review of Microbiology* 2006, **60**:107-130.
12. Taveirne ME, Sikes ML, Olson JW: **Molybdenum and tungsten in *Campylobacter jejuni*: their physiological role and identification of separate transporters regulated by a single ModE-like protein**. *Molecular Microbiology* 2009, **74**(3):758-771.
13. Park S: ***Campylobacter jejuni* Stress Responses During Survival in the Food Chain and Colonisation**. In: *Camplyobacter Molecular and Cellular Biology* Edited by Ketley J, Konkel M. Wymondham Horizon Bioscience 2005.
14. Harvey S, Lascelles J: **Respiratory systems and cytochromes in *Campylobacter fetus* subsp. *intestinalis***. *Journal of Bacteriology* 1980, **144**(3):917-922.
15. Sargent F, Berks BC, Palmer T: **Assembly of membrane-bound respiratory complexes by the Tat protein-transport system**. *Archives of Microbiology* 2002, **178**(2):77-84.
16. Schlindwein C, Giordano G, Santini CL, Mandrand MA: **Identification and expression of the *Escherichia coli fdhD* and *fdhE* genes, which are involved in the formation of respiratory formate dehydrogenase**. *Journal of Bacteriology* 1990, **172**(10):6112-6121.
17. Stewart V, Lin JT, Berg BL: **Genetic evidence that genes *fdhD* and *fdhE* do not control synthesis of formate dehydrogenase-N in *Escherichia coli* K-12**. *Journal of Bacteriology* 1991, **173**(14):4417-4423.
18. Hitchcock A, Hall SJ, Myers JD, Mulholland F, Jones MA, Kelly DJ: **Roles of the twin-arginine translocase and associated chaperones in the biogenesis of the electron transport chains of the human pathogen *Campylobacter jejuni***. *Microbiology* 2010, **156**(Pt 10):2994-3010.
19. Turner RJ, Papish AL, Sargent F: **Sequence analysis of bacterial redox enzyme maturation proteins (REMPs)**. *Canadian Journal of Microbiology* 2004, **50**(4):225-238.
20. Korolik V: **Aspartate chemosensory receptor signalling in *Campylobacter jejuni***. *Virulence* 2010, **1**(5):414-417.

21. Paster BJ, Gibbons RJ: **Chemotactic response to formate by *Campylobacter concisus* and its potential role in gingival colonization**. *Infection and Immunity* 1986, **52**(2):378-383.
22. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hackermuller J, Reinhardt R *et al*: **The primary transcriptome of the major human pathogen *Helicobacter pylori***. *Nature* 2010, **464**(7286):250-255.
23. Simon J, Gross R, Klimmek O, Kroger A: **The Genus Wollinella**. In: *The Prokaryotes: A Handbook on the Biology of Bacteria: Vol 7.* Edited by Martin Dworkin SF; 2006.
24. Bansal MP, Oborn CJ, Danielson KG, Medina D: **Evidence for two selenium-binding proteins distinct from glutathione peroxidase in mouse liver**. *Carcinogenesis* 1989, **10**(3):541-546.
25. Bansal MP, Cook RG, Danielson KG, Medina D: **A 14-kilodalton selenium-binding protein in mouse liver is fatty acid-binding protein**. *Journal of Biological Chemistry* 1989, **264**(23):13780-13784.
26. Ogasawara Y, Lacourciere GM, Ishii K, Stadtman TC: **Characterization of potential selenium-binding proteins in the selenophosphate synthetase system**. *PNAS* 2005, **102**(4):1012-1016.
27. Lacourciere GM, Levine RL, Stadtman TC: **Direct detection of potential selenium delivery proteins by using an *Escherichia coli* strain unable to incorporate selenium from selenite into proteins**. *PNAS* 2002, **99**(14):9150-9153.
28. Duret L, Gasteiger E, Perriere G: **LALNVIEW: a graphical viewer for pairwise sequence alignments**. *Compututional Applied Bioscience* 1996, **12**(6):507-510.
29. Lacourciere GM, Mihara H, Kurihara T, Esaki N, Stadtman TC: ***Escherichia coli* NifS-like proteins provide selenium in the pathway for the biosynthesis of selenophosphate**. *Journal of Biological Chemistry* 2000, **275**(31):23769-23773.
30. Jamba L, Nehru B, Bansal MP: **Redox modulation of selenium binding proteins by cadmium exposures in mice**. *Molecular and Cellular Biochemistry* 1997, **177**(1-2):169-175.

31. Meyer Y, Verdoucq L, Vignols F: **Plant thioredoxins and glutaredoxins: identity and putative roles**. *Trends Plant Science* 1999, **4**(10):388-394.

32. Flemetakis E, Agalou A, Kavroulakis N, Dimou M, Martsikovskaya A, Slater A, Spaink HP, Roussis A, Katinakis P: ***Lotus japonicus* gene Ljsbp is highly conserved among plants and animals and encodes a homologue to the mammalian selenium-binding proteins**. *Molecular Plant-Microbe Interactions* 2002, **15**(4):313-322.

33. Fonknechten N, Chaussonnerie S, Tricot S, Lajus A, Andreesen JR, Perchat N, Pelletier E, Gouyvenoux M, Barbe V, Salanoubat M *et al*: ***Clostridium sticklandii*, a specialist in amino acid degradation:revisiting its metabolism through its genome sequence**. *BMC Genomics* 2010, **11**:555.

34. Parrish JR, Yu J, Liu G, Hines JA, Chan JE, Mangiola BA, Zhang H, Pacifico S, Fotouhi F, DiRita VJ *et al*: **A proteome-wide protein interaction map for *Campylobacter jejuni***. *Genome Biology* 2007, **8**(7):R130.

35. Cobine P, Wickramasinghe WA, Harrison MD, Weber T, Solioz M, Dameron CT: **The *Enterococcus hirae* copper chaperone CopZ delivers copper(I) to the CopY repressor**. *FEBS Letters* 1999, **445**(1):27-30.

36. Katoh E, Hatta T, Shindo H, Ishii Y, Yamada H, Mizuno T, Yamazaki T: **High precision NMR structure of YhhP, a novel *Escherichia coli* protein implicated in cell division**. *Journal of Molecular Biology* 2000, **304**(2):219-229.

37. Pagni M, Ioannidis V, Cerutti L, Zahn-Zabal M, Jongeneel CV, Hau J, Martin O, Kuznetsov D, Falquet L: **MyHits: improvements to an interactive resource for analyzing protein sequences**. *Nucleic Acids Res* 2007, **35**(Web Server issue):W433-437.

Chapter Ten

Campylobacter and Selenoproteins: Concluding Summary and Future Directions

TABLE OF FIGURES

Chapter Ten - Campylobacter and Selenoproteins: Concluding Summary and Future Directions

The work presented in this PhD-thesis has described new features on *Campylobacter jejuni* biology and bacterial selenoprotein synthesis. Firstly, it has been shown that two previously uncharacterised genes of *C. jejuni* NCTC 11168 (Cj1500 and Cj1501) are necessary for the synthesis of a fully functioning formate dehydrogenase (FDH) enzyme, although the exact role of Cj1500-01 in this process is still not completely clear. Homologues of Cj1500 and Cj1501 also occur adjacently in other genomes throughout the bacterial kingdom, suggesting that the genes function in a coupled mechanism. It is conceivable that the proteins are performing the same role across phylogenetically diverse organisms. Secondly, identification of some of the genes involved in selenium incorporation and selenoprotein synthesis in *C. jejuni* NCTC 11168 have been identified. Selenoprotein synthesis in *Campylobacter* had not been reported upon previously, although the presence of a selenium-containing FDH in *C. jejuni* has been mentioned previously, albeit only indirectly within the literature [1-4]. This study has confirmed the roles of some of the genes predicted to be involved with selenocysteine synthesis (*selA* and *selB*) and the intracellular movement of selenium (Cj1500 and Cj1501) in *C. jejuni* NCTC 11168. There are several interesting and feasible directions in which to continue this investigation.

Insights into *C. jejuni* metabolism, gene regulation and selenoprotein synthesis

Explanation and clarification of any sequenced genome is an important field of research. The rapid advancements in DNA sequencing techniques now allow for rapid, easy and relatively cheap determination of genome sequences and, although this is a positive development, many genomes have been roughly, incorrectly or incompletely annotated. Numerous genes have been putatively annotated, and unless verified by biochemical experimentation, the annotation may remain wrong or incomplete. Furthermore, many previously annotated genomes are not curated or reannotated, and hence can influence

further annotation even if shown to be incorrect after release. This study has begun to explore the roles of some of the uncharacterised, or putatively annotated, genes in the Campylobacter genome. Investigations such as these are essential for the further understanding of bacterial genomics and genome function, and more specifically can increase our comprehension of an important food borne human pathogen, like *C. jejuni*.

This study has confirmed that *C. jejuni* NCTC 11168 has only one enzyme capable of metabolising formate, the FDH; Cj1511c-09c (*fdhABC*). The *C. jejuni* FDH is most likely a selenoprotein and the genes Cj1378 and Cj1379, the selenocysteine formation and incorporation genes, are required to produce a fully functioning FDH.

The *C. jejuni* NCTC 11168 genes Cj1500 and Cj1501 are also required for a fully functioning FDH and are likely to work in a coupled function. Currently, the favoured hypothesis is that they have a role in selenium transport and intracellular movement. Based on the results generated in this investigation, a speculation for the selenium up-take system is as follows: Cj1500 is acting as a transmembrane channel for the up-take of exogenous selenium and Cj1501, which is located in the cytoplasm, then has a role in guiding the selenium to the selenocysteine formation apparatus. When the Cj1500 and/or the Cj1501 are inactivated, the resulting *C. jejuni* phenotype is FDH deficient when examined using a benzyl viologen linked assay. However, this phenotype was partially restored by the addition of exogenous selenium. It is hypothesised that under high selenium conditions, selenium can enter the *C. jejuni* cells through an alternative, non-specific transporter and selenocysteine-containing FDH can be produced. This hypothesis was further confirmed by the very slow formate metabolism observed in the Cj1501 insertion inactivation mutant by ^{1}H NMR analysis of metabolites in the medium used to cultivate *C. jejuni*. This FDH activity was obviously below the level of detection of the benzyl viologen linked assay. If the selenium was entering the cell at a severely reduced rate by the action of an unspecific up-take system, this would explain the extremely low, but clearly detectable, FDH activity after overnight growth.

Future considerations for microarray design

This investigation has used a whole genome *C. jejuni* NCTC 11168 microarray for comparison of the transcriptomes of the wild-type and Cj1501 mutant. Although microarrays have been used for several years, the technology is constantly improved. This investigation has highlighted one particular area for microarray improvement; the need for inclusion of a probe pre-UGA codon within a selenoprotein coding gene.

The occurrence of selenoproteins within a genome can be predicted by the presence of the SECIS element down-stream of the UGA and by sequence homology to known selenoproteins. The *C. jejuni* NCTC 11168 microarray used in this investigation did not contain a probe up-stream of the UGA codon in the *fdhA* gene, although it did contain five probes downstream of the selenocysteine codon in *fdhA*. A clear down-regulation of the *fdhABC* transcription in *C. jejuni* NCTC 11168 (Cj1511c-09c) genes was observed in the Cj1501 insertion inactivation mutant, but there was no change in the Cj1514-13c genes which are on the same operon. Due to the absence of a probe in the *fdhA* segment upstream of the UGA selenocysteine codon, it could not be determined whether the down-regulation of *fdhA* transcription is throughout the gene or only downstream of the selenocysteine codon.

The simple addition, or redesigning, of the *fdhA* probes could have provided a clear answer to where exactly transcription of the *fdhA* is being down-regulated. Regarding whole genome microarray design, perhaps a future consideration should be the location of a probe before the selenocysteine codon, UGA, in a selenoprotein gene.

One of the possible hypotheses is that transcription of all FDH subunits takes place, however translation of the *fdhA* gene is most likely halted at the UGA codon due to a lack of available selenocysteine. This may cause the mRNA to destabilise, therefore translation of the FDH protein is halted at this point. The inhibition of translation may have been responsible for the apparent down-regulation of the FDH subunit genes when Cj1501 was inactivated by insertion of an antibiotic cassette (Figure 10.1). When bacterial cells are

grown without sufficient exogenous selenium, the formation of their selenoproteins can be ceased at the UGA codon [5].

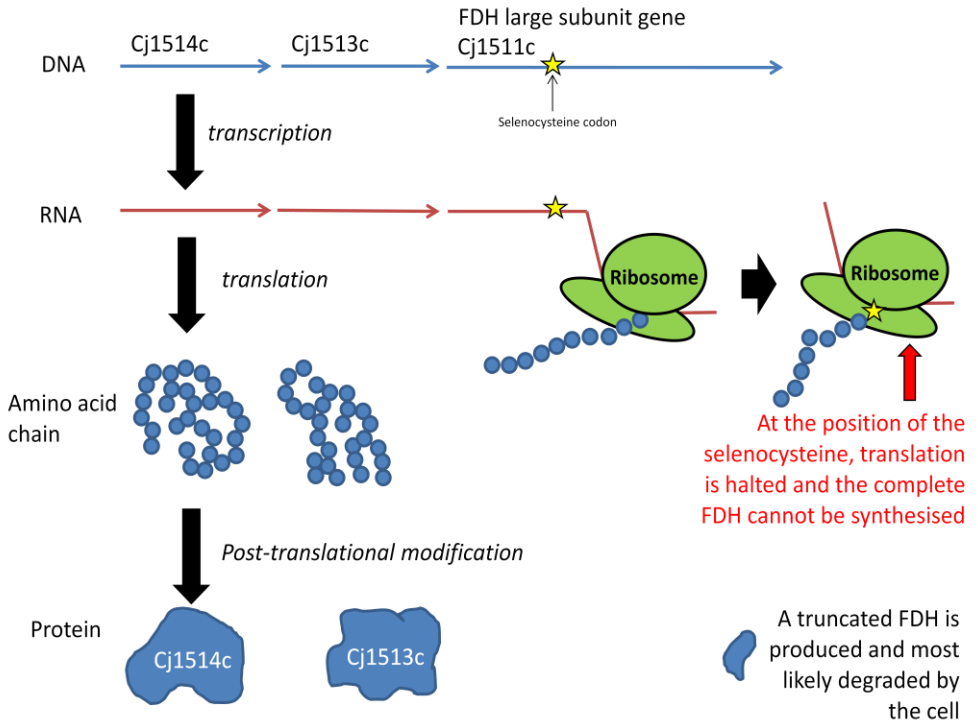


Figure 10.1 The possible transcription and translation of the FDH operon.
(This is the same figure as depicted in 8.21.)

The lack of a probe before the UGA is unfortunate, but the question of whether transcription is unaffected before the selenocysteine codon in the selenoprotein gene could be answered using reverse transcriptase PCR and primers specific for the area upstream of the UGA. For comparison, a pair of primers down-stream of the UGA and a pair flanking the UGA could also be used under the same experimental conditions. However, if the microarray had been designed with a probe before the UGA, this question could easily be answered. It is a noteworthy observation and the distribution of probes within a selenoprotein gene, or another unusual amino acid containing protein, should definitely be taken into consideration when designing future microarrays.

Selenoprotein formation in *C. jejuni*

This investigation has shown that *C. jejuni* NCTC 11168 contains a selenocysteine-containing FDH, and has established that *selA*; Cj1378 and *selB*; Cj1379 are most likely responsible for selenocysteine formation. This is the first example of a defined and biochemically proven role for the genes Cj1378 and Cj1379. The annotation of these genes was based on sequence homology to other previously sequenced microbial *selA* and *selB* genes.

Selenocysteine-containing FDHs are relatively common and this investigation has confirmed the FDH-activity and the selenium containing property, of a gene previously only annotated putatively as a selenocysteine-containing FDH. This information will add to the universal understanding of FDHs. Bacterial selenoprotein research, may be applicable in the understanding of other atypical proteins, where another unusual amino acid is inserted. Particular interest is given to the field of unusual codon usage, such as the UGA in *C. jejuni* coding for both a selenocysteine and as a stop codon, because it is intriguing that one codon can code for two different events in the same micro-organism.

In addition, this investigation has provided a predicted pathway for the genes involved in selenium cellular incorporation and usage in *C. jejuni* NCTC 11168; the pathway describes the entry of exogenous selenium into the *C. jejuni* cell, the formation of selenocysteine, the incorporation of selenocysteine into the FDH protein and the export of the FDH protein, via the TAT apparatus, to the periplasm. The pathway was predicted based on information available in the literature and the results generated during this investigation (Figure 10.2). The pathway for selenium incorporation and selenocysteine formation has never been reported before in Campylobacter.

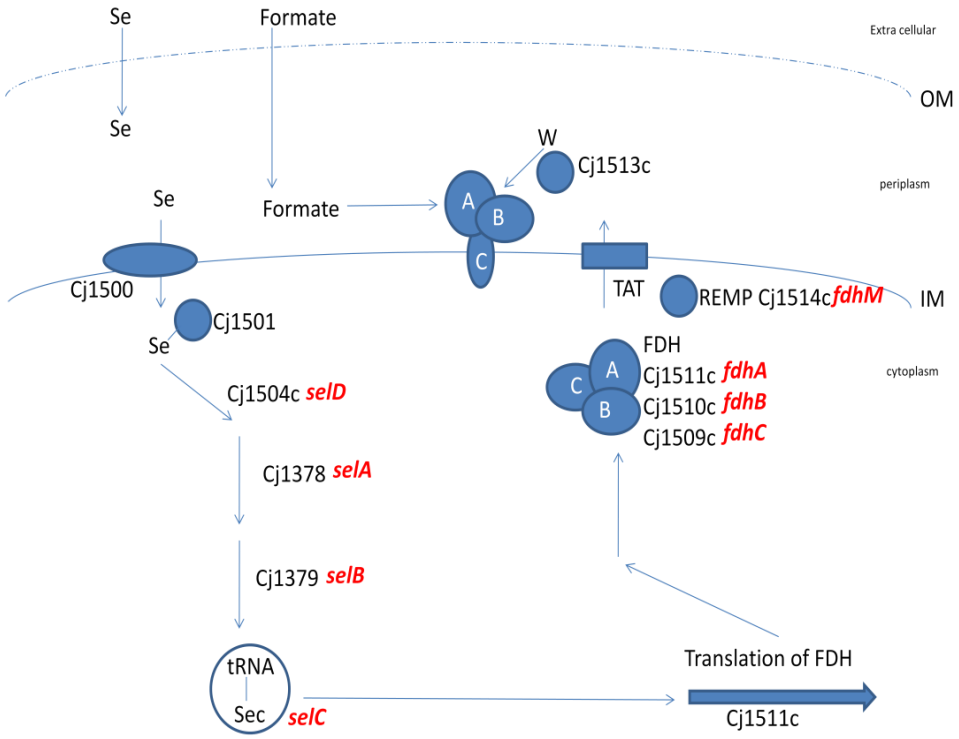


Figure 10.2 The possible function of the *C. jejuni* NCTC 11168 genes relating to the production of active FDH.
The gene names are shown in red. W – tungsten, Se – selenium, OM – outer membrane, IM – inner membrane, REMP – redox enzyme maturation protein, TAT – twin-arginine transporter, FDH – formate dehydrogenase. Note, this diagram is based on predictions formed after bioinformatic analysis, biological experimentation and the available literature (This is the same figure as depicted in 9.8)

Formate metabolism in *C. jejuni*

This investigation has used both an indirect (benzyl viologen linked enzyme assay) and a direct technique (NMR spectroscopy) to understand formate metabolism in *C. jejuni*. Weekaroon *et al.* discussed the importance of formate dehydrogenase in *C. jejuni* [6] and formate respiration in *C. jejuni* has been reported upon [7, 8]. This investigation has shown that formate is a substrate of *C. jejuni* respiration and is metabolised almost instantly from the growth medium. This suggests that formate is a preferred substrate of *C. jejuni*. In addition, it has been shown that *C. jejuni* NCTC 11168 only have one enzyme capable of formate metabolism. However, in contrast to this result, Mohammed *et al.* suggested that *C. jejuni* NCTC 11168 possesses several enzymes that are involved in the metabolism of formate through different metabolic pathways [8].

It has not been established whether FDH expression or activity is inducible by substrate or selenium in *C. jejuni*, and this would be an interesting area of research to pursue. It is known that *C. jejuni* are highly adaptable bacteria and able to survive by metabolising a range of bacterial by-products produced by the commensal gut bacteria. However, it is not known whether FDH is always expressed in *C. jejuni* cells and prepared to metabolise any available formate, or whether the enzyme is inducible by stimulation. Previous proteomic studies performed on *C. jejuni* grown in MEMα, a defined medium lacking formate, did not have detectable FDH protein subunits (Fran Mulholland, IFR, personal communication). This suggests that perhaps the expression of the enzyme is substrate or selenium inducible, but further testing would be required to back up this statement.

Future Research Directions

This investigation could continue in several directions. The possible future aspects of this work have been discussed in detail in the relevant summary sections of each chapter, however a collective list of some of the possible areas for expansion of this work are detailed briefly below:

- Further investigation of the Cj1501 insertion inactivation mutant to determine if it possesses a more robust acid-resistance phenotype than wild-type *C. jejuni* NCTC 11168.
- Crystal structure analysis of recombinantly purified Cj1501 protein to determine the function of the protein, and whether it has a recognisable selenium binding site.
- Mass spectroscopy of the recombinantly purified Cj1501 protein to determine whether it can bind a selenium ion.
- Microarray analysis of the Cj1501 mutant grown in a selenium supplemented medium to determine whether the down-regulation of the FDH genes observed without exogenous selenium is restored.
- Proteomic analysis of the Cj1501 mutant grown in selenium supplemented medium to support the conclusion that the FDH

(Cj1509-11c) protein complex is restored in the mutant in the addition of selenium.

- Proteomic analysis of wild-type *C. jejuni* and the Cj1501 insertion inactivation mutant grown in formate supplemented and formate free media to determine whether the FDH protein complex is inducible on substrate availability.
- Genetic complementation of the Cj1500-01 mutant with a non-native homologous gene pair from a species which is phylogenetically diverse from *C. jejuni* to determine whether the FDH phenotype can be restored, thus describing conserved gene function across species.
- Research into the bioavailability of selenium to micro-organisms, focussing specifically on the gastro-intestinal environments which *C. jejuni* inhabits.

Concluding Remarks

This investigation has furthered our understanding of the formation of selenoproteins in *C. jejuni*; which has never been reported previously. Figure 10.2 shows the hypothesised genes involved in the formation of FDH, including the selenium up-take system genes. The research in this thesis has confirmed specific roles to certain *C. jejuni* NCTC 11168 genes which were previously putatively annotated (Cj1378 and Cj1379). In addition, this project has started to elucidate the role of some previously uncharacterised genes which, prior to this investigation, had no predicted role based on the genome sequence (Cj1500 and Cj1501).

This work has provided interesting insight into a relatively unstudied field, and has significance not just within the Campylobacter research field, but the wider field of bacterial study. The conclusions formed in this investigation may be appropriate for application to several bacterial species, and may lead to the further understanding of more peculiar bacteria and uncommon amino acid containing proteins. This area of research has many exciting possibilities for future development. Studies such as this are essential because as the ability to sequence genomes becomes quicker, cheaper and faster, genomic data is being generated at an exponential rate, therefore the need to

understand these genomes is also increasing exponentially. Comprehension of bacterial genomes is necessary to further understand what information and function is stored within the genome of a single micro-organism, and how genomes are comparable between different species. Genome annotation relies heavily on the comparison of newly identified open reading frames with previously characterised genes. Therefore, it is essential that as much information as possible is gained using those sequences already readily available, and those micro-organisms which can be grown, genetically manipulated and experimented upon in the laboratory.

Chapter Ten – References

1. Hitchcock A, Hall SJ, Myers JD, Mulholland F, Jones MA, Kelly DJ: **Roles of the twin-arginine translocase and associated chaperones in the biogenesis of the electron transport chains of the human pathogen *Campylobacter jejuni***. *Microbiology* 2010, **156**(10):2994-3010.
2. Smart JP, Cliff MJ, Kelly DJ: **A role for tungsten in the biology of *Campylobacter jejuni*: tungstate stimulates formate dehydrogenase activity and is transported via an ultra-high affinity ABC system distinct from the molybdate transporter**. *Molecular Microbiology* 2009, **74**(3):742-757.
3. Taveirne ME, Sikes ML, Olson JW: **Molybdenum and tungsten in *Campylobacter jejuni*: their physiological role and identification of separate transporters regulated by a single ModE-like protein**. *Molecular Microbiology* 2009, **74**(3):758-771.
4. Stock T, Rother M: **Selenoproteins in Archaea and Gram-positive bacteria**. *Biochim Biophys Acta* 2009, **1790**(11):1520-1532.
5. Arner ES: **Recombinant expression of mammalian selenocysteine-containing thioredoxin reductase and other selenoproteins in *Escherichia coli***. *Methods in Enzymology* 2002, **347**:226-235.
6. Weerakoon DR, Borden NJ, Goodson CM, Grimes J, Olson JW: **The role of respiratory donor enzymes in *Campylobacter jejuni* host colonization and physiology**. *Microbial Pathogenesis* 2009, **47**(1):8-15.
7. Myers JD, Kelly DJ: **A sulphite respiration system in the chemoheterotrophic human pathogen *Campylobacter jejuni***. *Microbiology* 2005, **151**(1):233-242.
8. Mohammed KA, Miles RJ, Halablab MA: **The pattern and kinetics of substrate metabolism of *Campylobacter jejuni* and *Campylobacter coli***. *Letters of Applied Microbiology* 2004, **39**(3):261-266.

Chapter Eleven

Thesis Conclusions

Chapter Eleven - Thesis Conclusions

Continued elucidation of gene function is necessary and a hugely important part of modern biology. Understanding gene function, particularly in pathogens (or in model organisms which have similarity with known pathogens), is an important area for research. This knowledge can often be applied to drug design, pathogen eradication programs and the food safety industry.

Using robust biochemical and molecular biology techniques, the biological function of several previously uncharacterised or unannotated genes has been determined. In addition, using bioinformatic investigation, the roles of some further genes involved in possible biosynthetic pathways have been predicted.

A study such as this is essential as the number of sequenced and annotated genomes dramatically increases. Our ability to sequence and then annotate complete genomes has become quicker and cheaper, but occasionally mis-annotations occur. The existence of annotation errors in published genomes may have hindered the comparative genomics approach which has recently become so popular. Comparative genomics concerns the comparison of genomes to explain global genome evolution and plasticity, and to propose new functional annotation of genes on the basis of correlation or anti-correlation of gene occurrence, neighbourhood, fusion events and phylogeny [1]. Therefore; confirming, with biochemical techniques, the function of these putatively annotated genes is an important step in phylogenetic, phylogenomic and evolutionary studies.

This investigation has demonstrated the application and limitations of bioinformatic analysis using published genomes, protein sequences and various bioinformatic tools. Conclusions regarding the horizontal and vertical transfer of genes between species and the function of proteins have been drawn. Making predictions of gene function and origin based on bioinformatics alone has serious limitations. The exact function of a gene can only be indisputably determined with specific biochemical testing, such as recombination and expression of a gene in an alternative host, or inactivation of the gene followed by phenotypic investigation. However, whatever

conclusions are drawn, it is paramount that the correct control experiments have been performed simultaneously, so the conclusion cannot be described as erroneous. The speculation of gene evolution will always remain only as a prediction, because the hypotheses and theories can never be tested unequivocally. The assumptions of gene evolution are primarily based on the information available and the application of logical, consistent and plausible theories presented on the evolution of genomes.

Investigations such as this should be used in the continual curation and re-annotation of genome sequences. Efforts to annotate the genomes of a wide variety of model organisms are currently carried out by sequencing centers, model organism databases and academic/institutional laboratories around the world [2]. Standardisation of genome annotation has been discussed at length because different annotation methods and tools have been developed over time to meet the needs of biologists faced with the task of annotating biological data [2]. Some groups performing genome annotation work do not use literature references or annotate from literature to any great extent, but some, mainly prokaryotic genome annotation groups do. This is most likely because the prokaryotic species tend to have simpler gene structures and arrangements and smaller genomes. However, manual annotation using bioinformatic methods in conjunction with the available literature is more time consuming than the automated annotation methods. But, the ever growing evidence suggests that evidence-based functional annotation is the future of robust annotations, and therefore, investigations such as the one described in this thesis are essential in this era of readily available genome sequences.

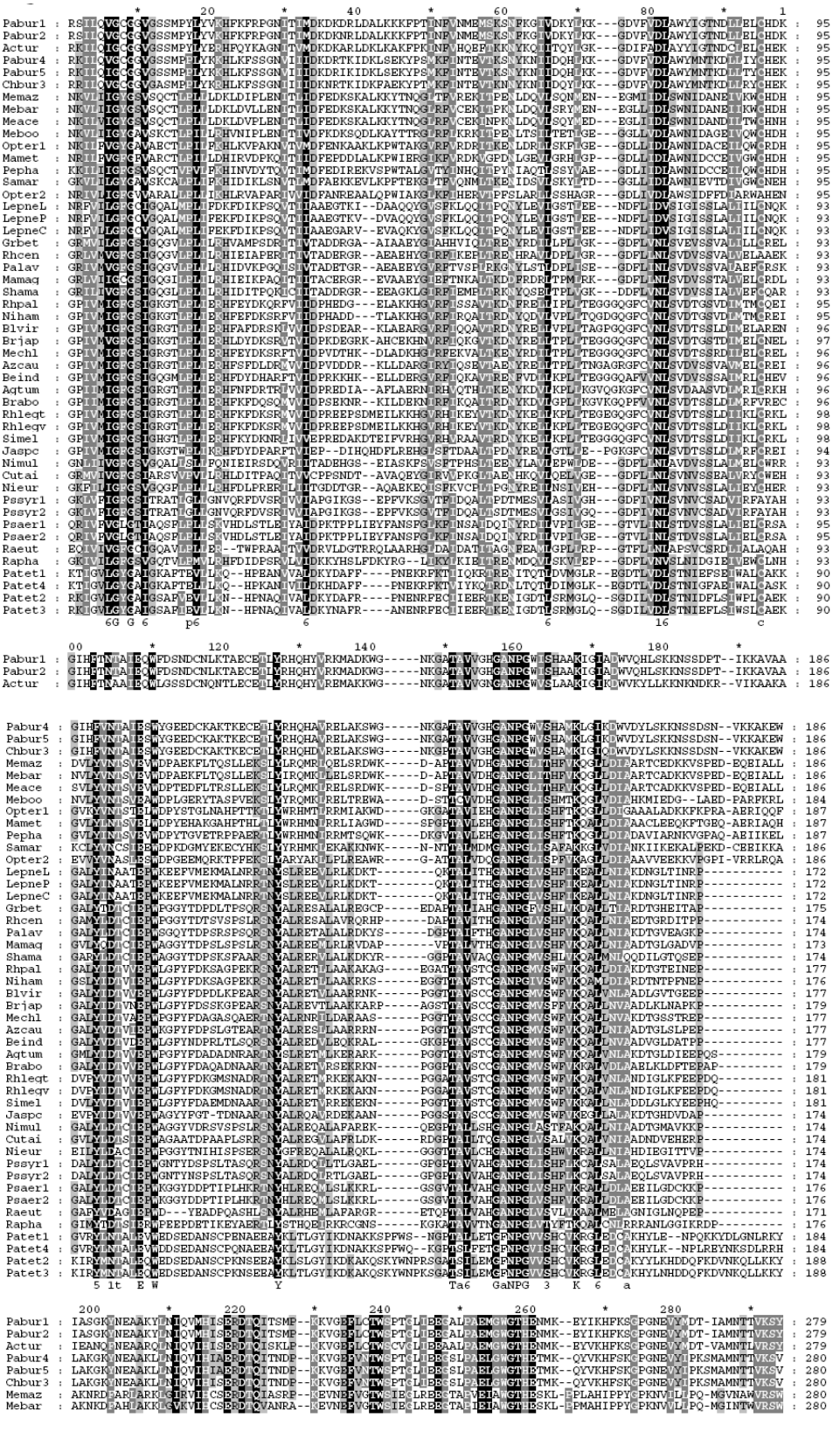
This thesis has determined the role of several either uncharacterised or unannotated genes across various bacterial species, both in polyamine biosynthesis and formate metabolism. The function of these genes has been shown using robust biochemical testing methods. This information can now be published and made available to the wider scientific community. The more genes which are characterised (rather than annotated based on sequence homology to other genes) the more robust the bioinformatic data is for future bioinformatic investigations.

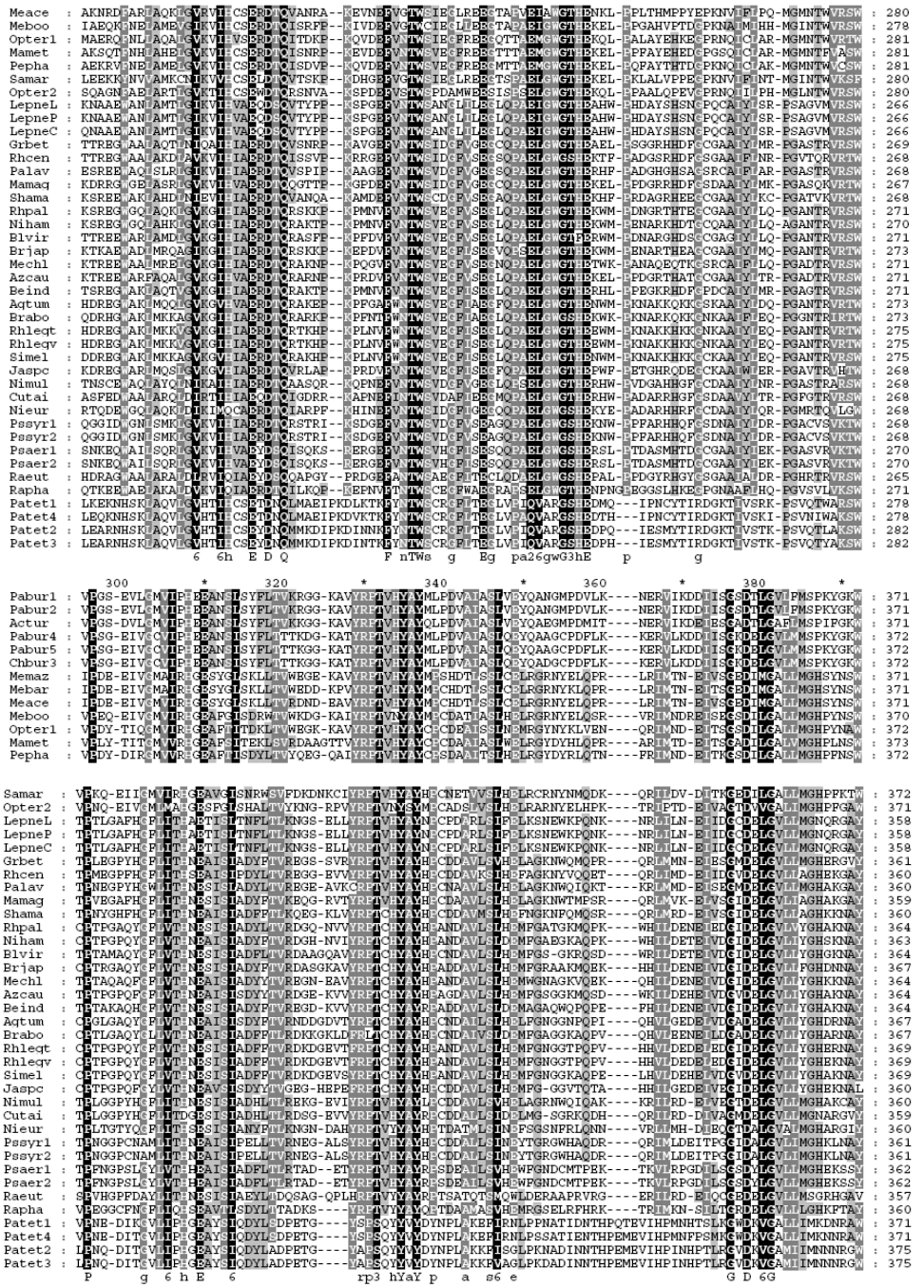
References – Chapter Eleven

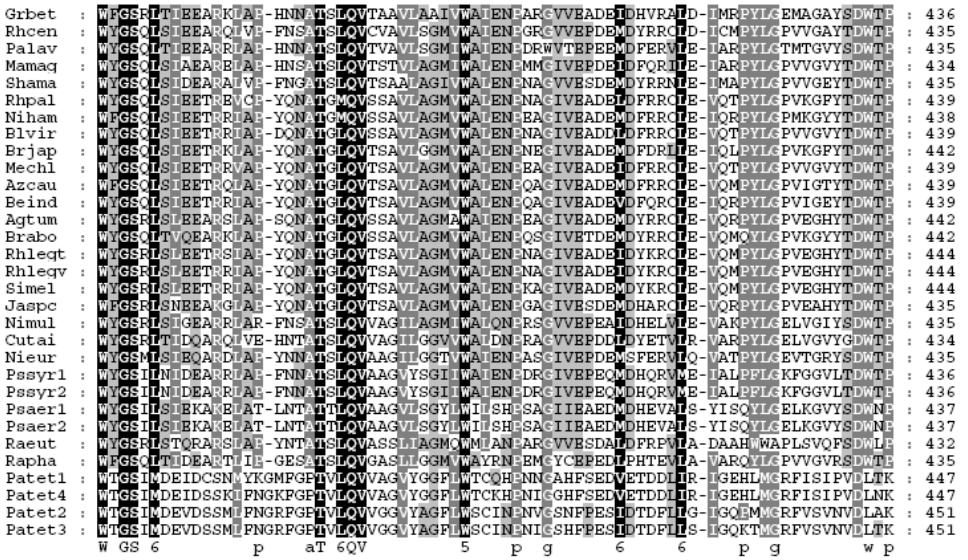
1. Armengaud J: **A perfect genome annotation is within reach with the proteomics and genomics alliance**. *Current Opinions in Microbiology* 2009, **12**(3):292-300.
2. Madupu R, Brinkac LM, Harrow J, Wilming LG, Bohme U, Lamesch P, Hannick LI: **Meeting report: a workshop on Best Practices in Genome Annotation**. *Oxford* 2010, **2010**:baq001.

Appendices

Appendix I The alignment on which the tree is based (Figure 4.5) Constructed by Tony Michael.







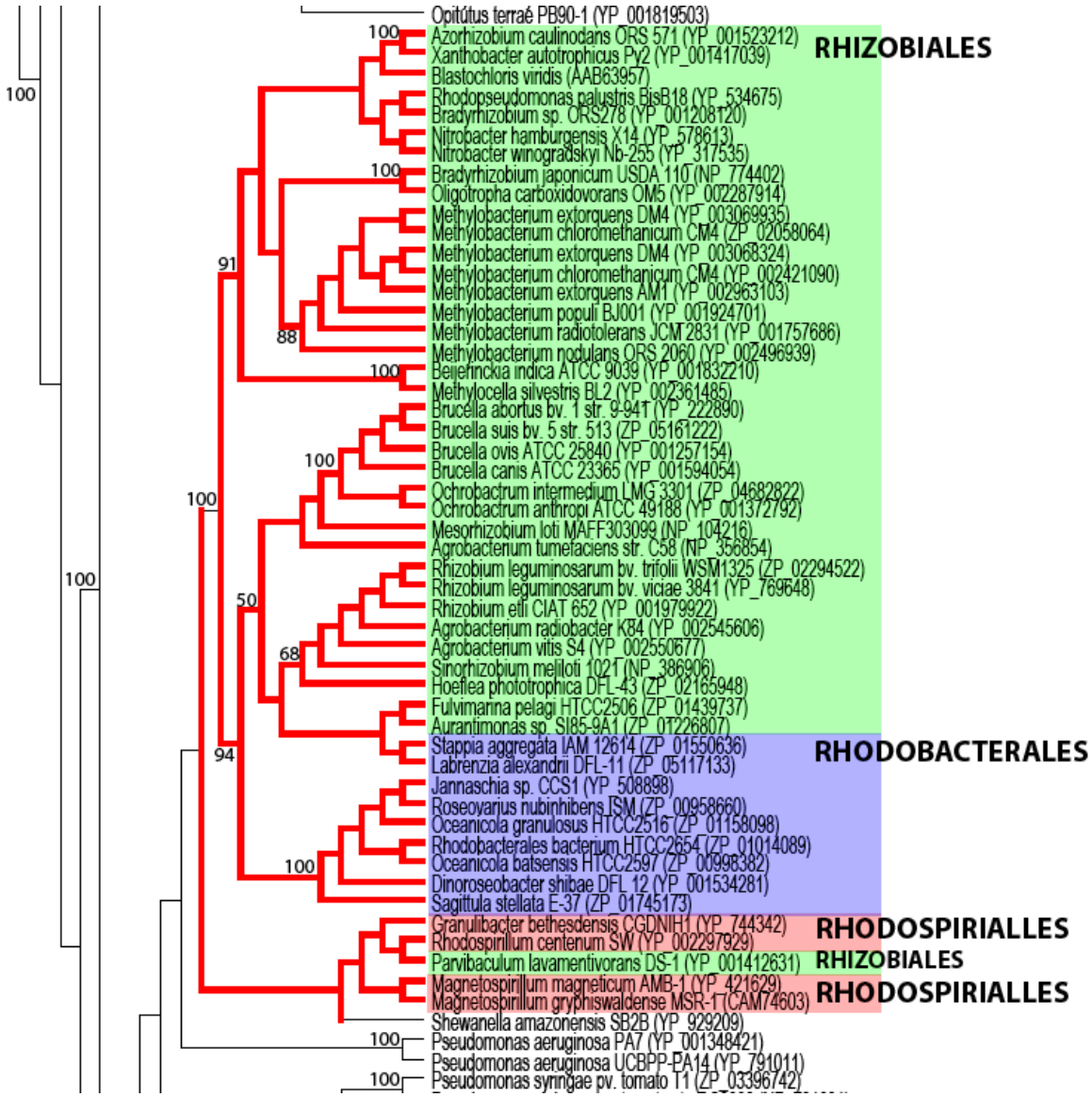
ClustalW alignment of HSS orthologues used for construction of the Neighbour joining phylogenetic tree presented in Figure 4.5

Pabur1, *Paramecium bursaria* Chlorella virus FR483 (YP_001425864);
Pabur2, *Paramecium bursaria* Chlorella virus MT325 (ABT13787);
Actur, *Acanthocystis turfacea* Chlorella virus 1 (YP_001427071);
Pabur4, *Paramecium bursaria* Chlorella virus NY2A (YP_001497501);
Pabur5, *Paramecium bursaria* Chlorella virus AR158 (YP_001498367);
Chbur3, *Paramecium bursaria* Chlorella virus 1 (NP_048585);
Memaz, *Methanosarcina mazei* Go1 (NP_632190);
Mebar, *Methanosarcina barkeri* str. Fusaro (YP_306169);
Meace, *Methanosarcina acetivorans* C2A(NP_616566);
Meboo, *Candidatus Methanoregula boonei* 6A8 (YP_001404208);
Samar, *Sawyeria marylandensis* (composite);
Opter1, *Opitutus terrae* PB90-1 (YP_001817583); Mamet, marine metagenome sequence (ECW30472);
Pepha, *Pelodictyon phaeoclathratiforme* BU-1 (YP_002017783);
Opter2, *Opitutus terrae* PB90-1 (YP_001819503);
Psaer1, *Pseudomonas aeruginosa* PA7 (YP_001348421);
Psaer2, *Pseudomonas aeruginosa* UCBPP-PA14 (YP_791011);
Grbet, *Granulibacter bethesdensis* CGDNIH1 (YP_744342);
Rhcen, *Rhodospirillum centenum* SW (YP_002297929);
Palav, *Parvibaculum lavamentivorans* DS-1 (YP_001412631);
Mamag, *Magnetospirillum magneticum* AMB-1(YP_421629);
Shama, *Shewanella amazonensis* SB2B (YP_929209);
Rhpa1, *Rhodopseudomonas palustris* BisB18 (YP_534675);
Niham, *Nitrobacter hamburgensis* X14 (YP_578613);
Blvir, *Blastochloris viridis* (AAB63957);
Brjap, *Bradyrhizobium japonicum* USDA 110 (NP_774402);
Mechl, *Methylobacterium chloromethanicum* CM4 (ZP_02058064);
Azcau, *Azorhizobium caulinodans* ORS 571 (YP_001523212);
Beind, *Beijerinckia indica* subsp. *indica* ATCC 9039 (YP_001832210);
Rhlegt, *Rhizobium leguminosarum* bv. *trifolii* WSM1325 (ZP_02294522);
Rhlegv, *Rhizobium leguminosarum* bv. *viciae* 3841 (YP_769648);
Simel, *Sinorhizobium meliloti* 1021 (NP_386906);
Agtum, *Agrobacteriumtumefaciens* str. C58 (NP_356854);

Brabo, *Brucella abortus* bv. 1 str. 9-941 (YP_222890);
Jaspc, *Jannaschia* sp. CCS1 (YP_508898);
Nimul, *Nitrosospira multiformis* ATCC 25196 (YP_412390);
Cutai, *Cupriavidus taiwanensis* (YP_002297929);
Nieur, *Nitrosomonas europaea* ATCC19718 (NP_841539);
LepneL, *Legionella pneumophila* str. Lens (YP_127746);
LepneP, *Legionella pneumophila* subsp. *pneumophila* str.Philadelphia 1 (YP_096502);
LepneC, *Legionella pneumophila* str. Corby (YP_001251255);
Pssyr1, *Pseudomonas syringae* pv. tomato T1 (ZP_03396742);
Pssyr2, *Pseudomonas syringae* pv. tomato str. DC3000 (NP_791691);
Raeut, *Ralstonia eutropha* H16 (YP_724563);
Rapha, *Ralstonia* phage RSL1 (YP_0019499984);
Patet1, *Paramecium tetaurelia* strain d4-2 (XP_001432408);
Patet4, *Paramecium tetaurelia* strain d4-2 (XP_001428569);
Patet2, *Paramecium tetaurelia* strain d4-2 (XP_001451562);
Patet3, *Paramecium tetaurelia* strain d4-2 (XP_001452042)

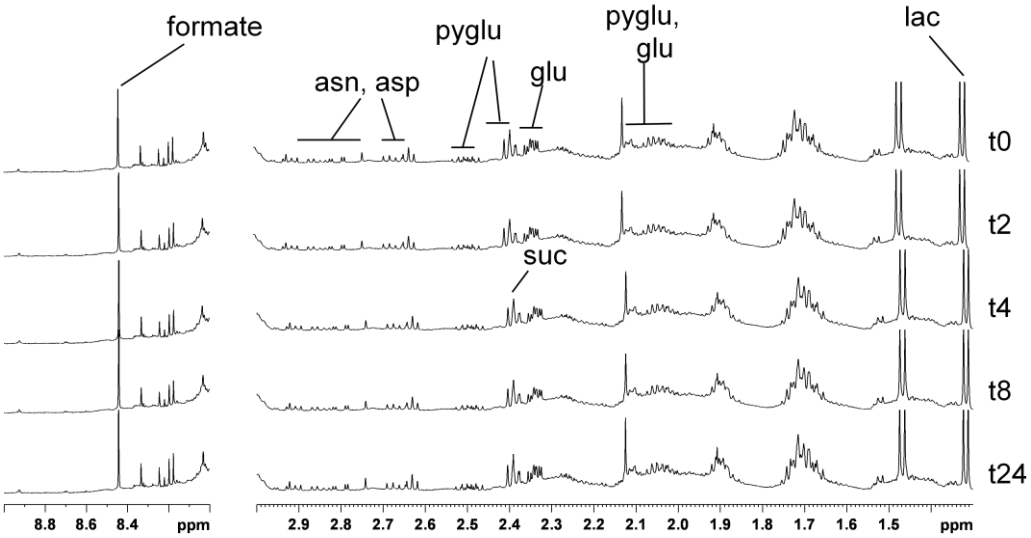
Appendix II Supplementary information for Figure 4.5 Neighbour joining tree of HSS with expanded α-Proteobacteria (red subtree)

Values represent percentage support from 1000 bootstrap replicates. Constructed by Tony Michael.

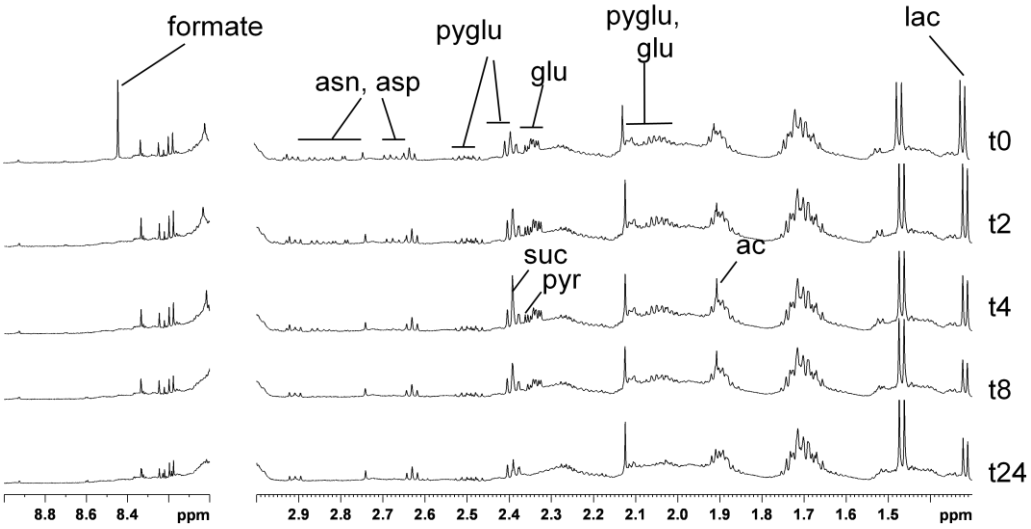


Appendix III Shows representative NMR trace data for the samples discussed in this section.

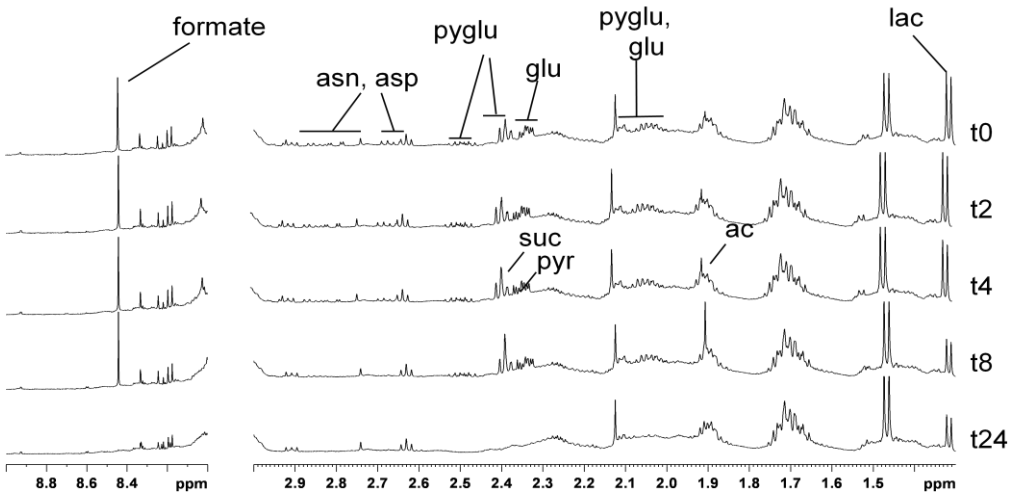
^{1}H NMR analysis performed by Gwen Le Gall.



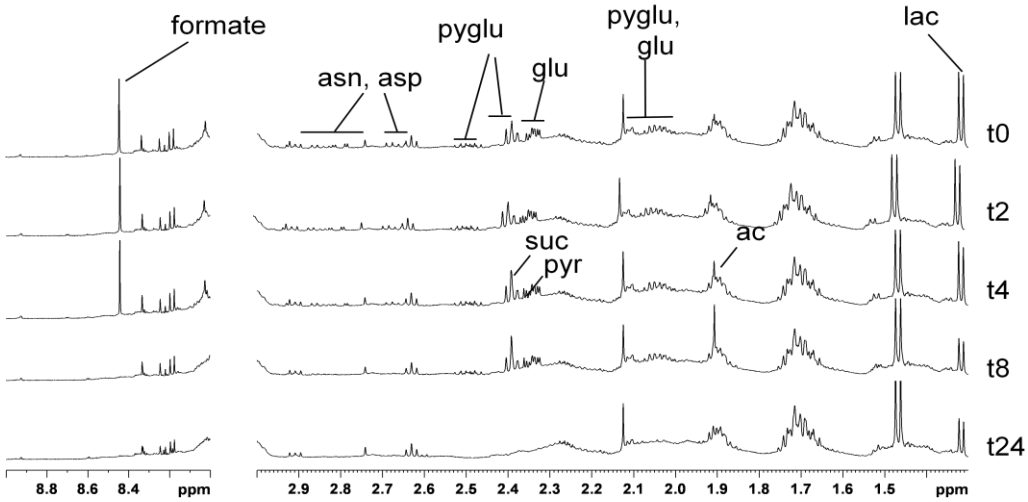
Overlay of ^{1}H NMR spectra of supernatant extracts from **control** cultures (Brucella broth only) reflecting metabolite profiles at 0, 2, 4, 8 and 24 hours; key to figure: lac, lactate; pyglu, pyroglutamic acid; glu, glutamic acid; suc, succinate; asp, aspartate; asn, asparagine.



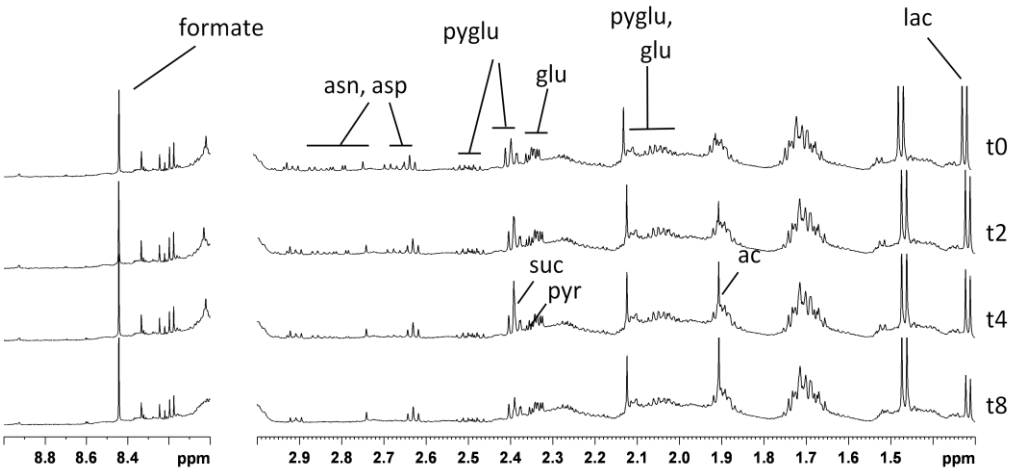
Overlay of ^{1}H NMR spectra of supernatant extracts from **wild-type** *C. jejuni* NCTC 11168 cultures reflecting metabolite profiles at 0, 2, 4, 8 and 24 hours; key to figure: ac, acetate; lac, lactate; pyglu, pyroglutamic acid; glu, glutamic acid; suc, succinate; pyr, pyruvate; asp, aspartate; asn, asparagine.



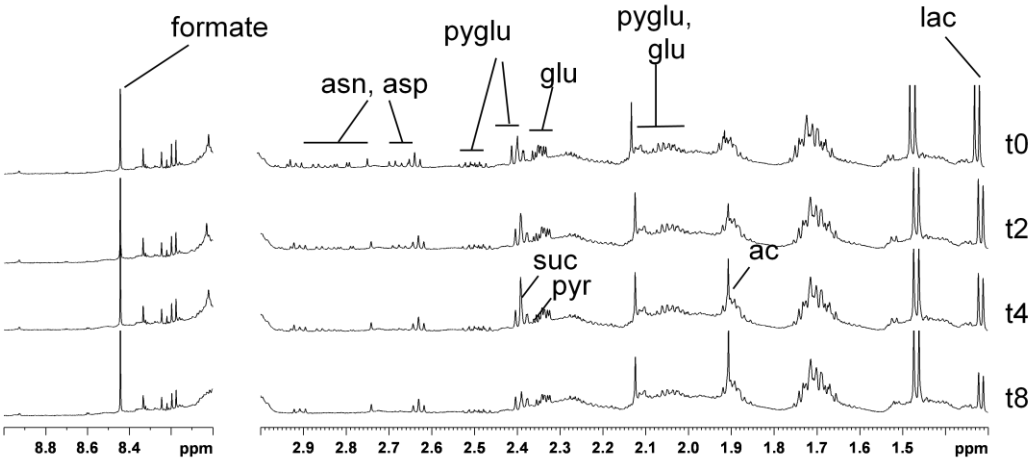
Overlay of ^{1}H NMR spectra of supernatant extracts from **Δ1501** *C. jejuni* cultures reflecting metabolite profiles at 0, 2, 4, 8 and 24 hours; key to figure: ac, acetate; lac, lactate; pyglu, pyroglutamic acid; glu, glutamic acid; suc, succinate; pyr, pyruvate; asp, aspartate; asn, asparagine.



Overlay of ^{1}H NMR spectra of supernatant extracts from **Δ1501 with 5 μM Selenium** *C. jejuni* cultures reflecting metabolite profiles at 0, 2, 4, 8 and 24 hours; key to figure: ac, acetate; lac, lactate; pyglu, pyroglutamic acid; glu, glutamic acid; suc, succinate; pyr, pyruvate; asp, aspartate; asn, asparagine.



Overlay of ^{1}H NMR spectra of supernatant extracts from **Δ1511** *C. jejuni* cultures reflecting metabolite profiles at 0, 2, 4, and 8 hours; key to figure: ac, acetate; lac, lactate; pyglu, pyroglutamic acid; glu, glutamic acid; suc, succinate; pyr, pyruvate; asp, aspartate; asn, asparagine.



Overlay of ^{1}H NMR spectra of supernatant extracts from **Δ1511 with 5 µM Selenium** *C. jejuni* cultures reflecting metabolite profiles at 0, 2, 4, and 8 hours; key to figure: ac, acetate; lac, lactate; pyglu, pyroglutamic acid; glu, glutamic acid; suc, succinate; pyr, pyruvate; asp, aspartate; asn, asparagine.

Appendix IV The probe positions and sequences for the *C. jejuni* NCTC 11168 microarray of gene Cj1511c *fdhA.*

The key is shown above the genetic code. There is not a probe upstream of the TGA codon (selenocysteine).

Cj1511c - *fdhA* subunit
MICRO-ARRAY PROBE SITES: ONE TWO THREE FOUR FIVE
TGA: Selenocysteine codon

ATGTCAAGTGTAGGTGAAAATATCAAGCTTACGCGTCGTTCTTTTCTTAAAATGGCAGCACTTTCAAGCC
TAGCAACACCGTTGCTTGCTAGAAGTGAAACTTTAAGAGAAGCGAGTGCTGATGAGCTTAAAGAAGCTTA
TGAAGGTAGTAAAAAAGTAAAAACAGTTTGTACAGCCTGTTCTGTAGGCTGTGGGATTATTGCAGAAGTG
CAAAATGGTGTTTGGGTGCGTCAAGAAATCGCTCAAGATCATCCTGTGAGTTCAGGTGGACATTGCTGCA
AGGGATCTGATATGATCGATATGGTGCGTTCACATGTGCGTTTAAAATATCCTATGAAAAAAGAAAATGG
AGAATGGAAACGCATAAGTTATGAGCAAGCTTTAAGTGAGATCGGAGAAAAACTAGCTGCTTATCGTAAA
GAAAATCCTGAAAGTGTTATGTTTTTAGGTTCTGCAAAACTTAATAATGAACAAGCTTATTATATAAGAA
AATTTGCAGCATTTTTTGGAACGAACAATGTAGATCATCAAGCTAGAATTTGACACAGCGCAACAGTCGC
CGGTGTGGCGAATACATTTGGTTATGGCGCTATGACAAACCATCTTGGAGACATCCAAAGAAGTAAATGT
ATCATTATCATTGGAGCAAATCCAGCGGTAAATCACCCTGTGGGTTTTAGACATTTCTTAAAAGCAAAAG
AAAAAGGTGCAAAGCTTATCGTTGTAGATCCTAGATTTACAAAAAGTGCAGCAAAAGCGGATATTTATGC
AAGAATTCGTCCAGGAACTGACATTGCTTTCATGTATGGAATGTTAAAAATCATTTTTGATGAAGGCTTG
GAAGATACAAAATATCTTGATGAAAGAGTTTTTGGAATTGATAAAATTCGTGAAGAAGCAGCAAAATGGA
CAGTTGAAGAAGTGGAAAATGTTACAGGGATCTCTAAAGAACTTTTGGTGCAAATCACTCACGAAGTAGC
TAAAAACAAACCAACTACGCTTATTTGGGCTATGGGTTTAACTCAACATACTGTAGGAACTTCAAATACA
CGTTTAGCACCTATTGTGCAAATGGTACTTGGAAATATAGGTAAATTTGGTGGTGGAGTTAATATCTTAC
GCGGACACGATAATGTTCAAGGTGCTTCAGATATGGCTTGTTTGAGCGAAAATTTACCAGGTTATTATCC
TTTAAATGAAGCGACTTGGAGATACTATGCCAAAATTTGGGGTGTGGATTATGAGTGGCTTTTAGGAAAT
TTTGTAAGCAAAGATTGGATGCATAAAACTGGGCTTTCACTTGCTAGATGGTGGGCAGCGGCTTTAAATG
GAAAAGATGGAAATGATGCTATCGATAATGCAGGAACGCCTTTAAAAGCTTTAGTAGTTATGGGAAATGG
TATTACTTCGACTGCACAACAAGTAAAAGTTAAAGAAGGTTTGGAAGCTTTAGAGCTTTTAGTTTTAGCC
GATCCTTTTGTAAATGAAGCAGGGATTATTGCAGAAAGAAAAGATGGAATTTATCTTTTACCTGCAGCAA
CACAGTTTGAAACAAGCGGAAGTGTAACAGCAACAAATCGTAGCGGACAATGGAGATTTAAAGTTGTAGA
TCCACTTTATGAAAGCATGGAAGATCAAGAAATTTTATTTGAGCTTGCTAAAAAATTAGGTTTTTATGAA
GACTTTACCAAGACTTTACGCGATGAAAAAGGTGAGATTGTTTGGCCTGAAAATGCAACAAGAGAAATTG
CAAAAGCAGTTAGAAGTATAGGGCTTAATGGTTGGAGTCCTGAAAGGCTTAAAAGCATACTTTATACTG
GGATAAATTTGATGAGGTAACTTTAGAAGGAAAAGACGAAGTTGCTGGCGAGTATTACGGGCTTCCTTGG
CCTTGTTGGAGTGATAAGCACCCAGGTTCTCCTGTGCTTTATAATACCGACATTGAAGTAGCAAAAGGTG
GTATGGGCTTTAGAAACAATTTTGGACTTGAATATGAAGGAGAGAGTTTACTTGCTAAGAATGCACCTTT
AAATTCACCTATTGATACAGGTTATCCACAAATTACTAAAGATAATATAGAAAAAGTTTTAGGTATTACT
TTAAGTGCTCAAGAAAAAGAAAAAATGGGATCAACTTGGTCTTATGATGATAGCAATATCATAGCAACTA
AATGTATAGAAAAAGGCATAGTTCCTTATGGAAATGCTAAAGCTAGAGCTGTAGTTTGGACTTTTAAAGA
TAAAATTCCACTCCACCGTGAGCCTTTACATTCTCCAAGAAACGATTTGGTGCAAAAATATCCAAGTTTT
GAAGATCAAAAAGCGCTTTACCGCGTGGATACTAAATTTGTTTCTGTGCAACAAGCTAAGGATTATTCTA
AAGAATTCCCGCTTAATCTTGTTACTGCAAGACTTGTGAATTTAAATGGTGCAGGTATGGAAAATAGAGC
ATCTATGTATCTAACGCGTTTAACACCTGAAATGTTTTGTGAGATCAATCCTGAATTAGCAAAAGAGCAA
GACATTAAAGCAGGAGATATGATTTGGGTGCATTCTCCAGAAGGTACTAAAATTCATGTAAGAGTGAAGG
TTAATCCAGGTGTAGCAAAAGATATGATTTTCTTGCCTTTCCATTTTACAGGAGTTATGCAAGGCGTGGA
TTTAACGCACAATTTCCCAGAAGGAACTAAGCCTTATGCAAGCGGAGAAAGTGCTAATACGGTAACAAAT
TATGGTTATGATATCATGTGTCAGATTCCAGAGACTAAAGGTGGACTTTGTAGAATCAGTAAGGATGGAA
AATGA

Appendix V The results of SIM (program which finds a user-defined number of best non-intersecting alignments between two protein sequences or within a sequence) using *C. jejuni* Cj1501 (C) and *E. coli* YDFZ (E)
using the parameters; Comparison matrix: BLOSUM62, Number of alignments computed: 20, Gap open penalty: 12, Gap extension penalty: 4

```
 54.5% identity in 11 residues overlap; Score: 29.0; Gap frequency: 0.0%

 E             27 KILSIDTEGLT
 C             57 KVLEIDQSGPT
                  * * **  * *

 ---------------------------------------------------------------------------

 42.9% identity in 7 residues overlap; Score: 19.0; Gap frequency: 0.0%

 E             25 TGKILSI
 C             30 SGEILEV
                   * **

 ---------------------------------------------------------------------------

 42.9% identity in 7 residues overlap; Score: 18.0; Gap frequency: 0.0%

 E             56 LAPLDLI
 C             19 IATLDVL
                   * **

 ---------------------------------------------------------------------------

 60.0% identity in 5 residues overlap; Score: 16.0; Gap frequency: 0.0%

 E              3 TYDRN
 C              4 TYSLN
                  **  *

 ---------------------------------------------------------------------------

 66.7% identity in 3 residues overlap; Score: 14.0; Gap frequency: 0.0%

 E             50 EGC
 C             12 EAC
                  * *

33.3% identity in 6 residues overlap; Score: 14.0; Gap frequency: 0.0%

E             30 SIDTEG
C              6 SLNLQG
                 *    *

---------------------------------------------------------------------------

27.3% identity in 11 residues overlap; Score: 14.0; Gap frequency: 0.0%

E              8 RNAITTGSRVM
C             49 QDAKNRGFKVL
                   *   *  *

---------------------------------------------------------------------------

27.3% identity in 11 residues overlap; Score: 13.0; Gap frequency: 0.0%

E             19 VSGTGHTGKIL
C             61 IDQSGPTLRFL
                     * *   *

---------------------------------------------------------------------------

66.7% identity in 3 residues overlap; Score: 13.0; Gap frequency: 0.0%

E              6 RNR
C             52 KNR
                  **

---------------------------------------------------------------------------

20.0% identity in 10 residues overlap; Score: 13.0; Gap frequency: 0.0%

E             40 QIRRGKTVVV
C             27 KLKSGEILEV
                    *    *

---------------------------------------------------------------------------

75.0% identity in 4 residues overlap; Score: 13.0; Gap frequency: 0.0%

E             10 AITT
C             18 AIAT
                 ** *
```

Appendix VI An alignment of *Carboxydothermus hydrogenoformans*, *Nitrosomonas europaea* and *Campylobacter jejuni* Cj1500-01 homologues as identified by String T-COFFEE_distribution_Version_8.93, output = clustalw,msf,phylip,score_html, ncore=4, outorder=aligned.
The sequences were made by taking the two protein sequences (from the homologues of Cj1500 and Cj1501) and combining them without an adjoining space. Therefore, the last 75 amino acids represent the Cj1501 homologous region. The alignment is colored according to its consistency with all the multiple sequence alignment used to compute it. The colour represents the reliability of the alignment. Regions in red have a high consistency and are more likely to be correctly aligned. Regions in green/blue have the lowest consistency.

```
Campylobacter     ----LNSFKQKYLINFWDNSRSMIALGILSAVYFGIFGGVWAVTGEMTRW
Carboxydothermus  MGEIIKSYYQKVIAEFWDQKTAVVLLGILSGLYFGTVGVVWAVTGEFTRW
Nitrosomonas      MS--LAEFRAQYLVRFWSPIPALLALGVASAYYFAITGTFWAVTGEFTRW
                      : .:  : : .**.   ::: **: *. **.  * .******:***

Campylobacter     GGEFLELLGMNLDGYSYYQKQNLNGTPLTRTDGIMLIGMFIGCLVAALLA
Carboxydothermus  GASFLKLIGVDLSPYTYLKIIKYKGTVLTRIDGVMVLGMFAGALIAALFG
Nitrosomonas      GGHIAALLGFSPQQWSYFQLIGLNGSPLERIDGVMIIGMFAGALCAALWA
                  *. :  *:*.. . ::* :    :*: * * **:*::*** *.* *** .

Campylobacter     NKVKFRLPASNIRIFQAIVGGILSGYGARLAFGCNLANFFTGLPYFSLHT
Carboxydothermus  QNFKLRIPTAK-RVLQALIGGIIAGFGTRLAMGCNLAALFTGIPQFSLHT
Nitrosomonas      GNVQLRWPTSRRRLAQGLIGGIIAGFGARLAMGCNLAAFFTGIPMFSLHA
                   :.::* *::. *: *.::***::*:*:***:***** :***:* ****:

Campylobacter     WLFTVFMVLGIYLGVKICNTSFFKPKAKLERVNKENLPLNKQSL--RTKL
Carboxydothermus  WFFTLGTIFGTYIGIKITLSPYFRGEPKLVKASEF----NAGNLKANTQI
Nitrosomonas      WAFMLTTVIGAWIGVKLCLLPFLRTPLRLDTAPSS-LFADTASLARRARL
                  * * :  ::* ::*:*:   .:::   :*  . .     :  .*  .:::

Campylobacter     YFNLGILLFIAFLVWVFYLVFTNGNISTQNKQSLLALALIFGFVFGFVIS
Carboxydothermus  QPFLG---------WLGLIVFAGILLSRPEMPNNLKLATVFGFAFGFLIQ
Nitrosomonas      QNRLG---------LLIAVLVLGFAAWRFETSLVLGLAVLFGVFFGAVIE
                     **          :  ::. .      :    * ** :**. ** :*.

Campylobacter     RGQICFTSCFRDLFLFGRDNAIKGALIGMIIASLIAFAFILQGHTSKLIE
Carboxydothermus  KGQVCFTSAFRDLWLVGRTTTLKALVWGMAVQMLLTAAFIAKGTPAKVLW
Nitrosomonas      RGQICFTSAARDLWTTGRTRIAYGILLGMVVACLGTFGAIALGATPKIFW
                  :**:****. ***:  **     . : ** :  * : . *  * ..*::

Campylobacter     LSPAVAVGAFLFGFGIVFAGGCECGWTYRAFEGQSHFIIVGIANIIGTMI
Carboxydothermus  AGPNALIGGLLFGIGIVIAGGCETGWMYRSMEGQVHFWFVGLGNVIGATI
Nitrosomonas      MGPNAALGGLLFGIGIVLAGGCETGWMYRAMEGQVHFWIVGIGNVIGGTL
                   .* . :*.:***:***:***** ** **::*** ** :**:.*:**  :

Campylobacter     LALSYNF-LPKAFKE-GIKINLLTEFGNLNGFFINL--ILFILMFVFVVF
Carboxydothermus  LFLVWDKGVYKYLAEPFPKFSLIEHFGYLPAFILTG--LFLLGLYLWADL
Nitrosomonas      VAIFWDE-LGGTLALPYPKINLLEYLGAGTGLLLSLAGLMLAMLLVYLN-
                  : : ::  :   :     *:.*:  :*   .:::.   :::  : ::

Campylobacter     ---YKKHFFKNQLKGV-----KITYSLNLQGEACPYPAIATLDVLPKLKS
Carboxydothermus  REASGRSMIKNKTMSVRRVMEEKRYVLDLRGEPCPYPVVYSLQVLAELES
Nitrosomonas      ----ARRFAVREGLARMNHNHQPDLSLDLRGEHCPYNAIATLEALADMTA
                       : :  .:  .      :    *:*:** *** .: :*:.*..: :

Campylobacter     GEILEVLCDCPQSINSIPQDAKNRGFKVLEIDQ-SGPTLRFLIQK-P
Carboxydothermus  GALLEILADCPQSFKSVPEEVVKAGYEMVEPPQKIGPTLRFLVRK-P
Nitrosomonas      GQVLEVITDCAQSVNGIPEDARAKGYDCLAVEQ-HGPLFRFLIRVPG
                  * :**:: **.**.:.:*::.   *:. :   *  ** :***::
```